The Epoch of Reionization

Marius Berge Eide

München 2021

The Epoch of Reionization

Marius Berge Eide

Dissertation
der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Marius Berge Eide
aus Oslo

München, den 30.10.2020

Erstgutachter: Prof. Dr. Simon D. M. White
Zweitgutachter: Prof. Dr. Gerhard Börner
Tag der mündlichen Prüfung: 17. Dezember 2020

Contents

List of Figures

List of Tables

Zusammenfassung

Die Epochen der kosmischen Erwärmung und Reionisierung sind Teil des Übergangs in dem das dunkle Zeitalter des Universums mit einer kosmischen Morgenröte endete. Irgendwann während der ersten Milliarde Jahre bildeten sich die ersten Galaxien und mit ihnen die ersten Lichtquellen, deren Strahlung das intergalaktische Gas erhitzte und ionisierte. Die Details dieses Phasenübergangs sind das Thema dieser Arbeit.

In Erwartung entscheidender Beobachtungen des frühen Universums mit Radioteleskopen haben wir uns die Aufgabe gestellt, die kosmische Morgenröte in einer neuartigen Weise zu modellieren und zu simulieren. Wir haben dafür ein System zusammengestellt, in dem wir theoretische Modelle von Hitze-, Ionisierungs- und Lichtquellen mit einer Simulation der Entstehung der kosmischen Struktur (und damit der Galaxien) koppeln und somit zusammen in einer hochpräzisen Strahlungstransportsimulation verwenden können. Dies erlaubt uns Vorhersagen über den Erwärmungs- und Ionisierungsprozess.

Wir nehmen unter anderem an, dass Sterne, Röntgendoppelsternsysteme und das interstellare Medium ionisierende Strahlung aussenden, die aus allen Galaxien austritt, und dass die massereichsten Galaxien auch massereiche Schwarze Löcher beherbergen, deren Akkretionsscheiben Quellen für heizende Strahlung sind. Um die Verteilung und die Strahlungseigenschaften der Quellen vorherzusagen, verwenden wir die hydrodynamische Simulation MassiveBlack-II mit ihrem Volumen von $(100\ \mathrm{cMpc}\ h^{-1})^3$ als unsere kosmische Modellumgebung. Wir verfolgen die Strahlung mit CRASH, einem multifrequenten, ionisierenden, dreidimensionalen kosmischen Strahlungstransferprogramm für die Weiterverarbeitung von hydrodynamischen Simulationen.

Unsere Haupterkenntnis ist, dass Sterne wahrscheinlich der Haupttreiber der Reionisierung von intergalaktischem Wasserstoff sind. Schwarze Löcher sind die einzigen Quellen, die Helium vollständig ionisieren können, weshalb wir erwarten, dass sie die Helium-Reionisierung zu späteren Zeiten dominieren. Die Gastemperatur wird von ihrem Hauptionisator, den Sternen, bestimmt, kann aber um Zehntausende von Grad weiter erhöht werden, wenn sich ein Schwarzes Loch in der Nähe befindet. Das interstellare Medium und die Röntgendoppelsterne treiben die Reionisierung nicht an, erwärmen und ionisieren aber dennoch das neutrale intergalaktische Gas geringfügig - ausreichend, um deutliche Signaturen zu hinterlassen, die mit gegenwärtigen und zukünftigen Radioteleskopen beobachtet werden können.

Wir untersuchen auch die seit langem bestehende Frage nach der Rolle der Schwarzen Löcher. In der hydrodynamischen Simulation beherbergen nicht alle Galaxien ein solches,

aber mithilfe eines neuronalen Netzwerks, trainiert auf denjenigen Galaxien, die eines beherbergen, bestücken wir künstlich alle Galaxien mit Schwarzen Löchern. Wir sagen eine Population von schwachen Schwarzen Löchern voraus, die einen signifikanten, aber nicht dominanten Beitrag zur Wasserstoff-Reionisierung leisten könnten. Zukünftige Raytracing-Simulationen sind erforderlich, um ihre thermischen Signaturen vorherzusagen.

Wir schließen diese Arbeit mit dem Aufruf, zukünftige Radiobeobachtungen vom Mond aus durchzuführen. Unser Zweig der Astronomie hat das Potential, die Menschheit in eine neue Ära der Zusammenarbeit und der friedlichen Erforschung des Weltraums zu führen.

Summary

The Epochs of Cosmic Heating and Reionization are the early times in which the Dark Ages of the Universe concluded with a Cosmic Dawn. Sometime during the first billion years, the first galaxies formed, and with them came the first sources of light. Their radiation heated and ionized the gas that spanned the space between galaxies. The details surrounding this phase transition is the topic of this thesis.

Awaiting decisive radio telescope observations of the early Universe, we have undertaken the task of modelling and simulating Cosmic Dawn in an unprecedented manner. We assembled a framework in which we combined theoretical models of sources of heat, ionization and light with a simulation of cosmic structure (and hence galaxy) formation and used these in a high-precision radiative transfer simulation. This allowed us to predict the heating and reionization process.

We assumed that stars, X-ray binary systems and the interstellar medium emit ionizing radiation that escape from all galaxies, and that the most massive galaxies also host massive black holes, whose accretion disks are sources of warming light. To predict the location and radiative properties of the sources, we used the $(100$ cMpc $h^{-1})^3$ hydrodynamic simulation MassiveBlack-II as our cosmic environment. We traced the radiation with the multifrequency ionizing three-dimensional 'Cosmic RAdiative transfer Scheme for Hydrodynamics', CRASH.

Our main finding is that stars are likely to be the main driver of reionization of hydrogen gas. Black holes are the only sources that may fully ionize helium, thus we expect them to dominate helium reionization at later times. The gas temperature is determined by their main ionizer, the stars, but may be raised tens of thousands of degrees further if a black hole is in the vicinity. The interstellar medium and the X-ray binaries do not drive reionization, but nevertheless heat and ionize neutral intergalactic gas ever so slightly—sufficiently to leave distinct signatures to be observed with present and future radio telescopes.

We also investigate the long-standing question of the role of black holes. Not all galaxies in the hydrodynamic simulation host one. Training a neural network on those that did, we mock black holes and place them into all galaxies. We predict a population of faint BHs that may provide a significant, but not dominant contribution to hydrogen reionization. Future ray tracing simulations are needed to predict their thermal signatures.

We conclude this thesis with a call for future radio observations to be done from the Moon. Our branch of astronomy has the potential to usher humanity into a new era of collaboration and peaceful space exploration.

Chapter 1

The Beginning

This chapter is dedicated to introducing and motivating the work that is presented throughout the remainder of this thesis. We will construct a vantage point from which we will overlook the themes we explore in a rapidly changing astronomical landscape. The overarching theme of this thesis can be put as two simple, but fundamental questions:

1. What was the temperature of the early Universe?

2. How did the Universe become ionized?

Although they are simple, they are by no means trivial. We will consider the cosmological epoch that began roughly a few hundred thousand years after the beginning of times, and which lasted for approximately one billion years. The timing of the beginning, duration, and conclusion of this epoch follow as answers to our questions. This cosmic saga begins with the *Dark Ages*, which transitioned through an *Epoch of Heating* during the *Cosmic Dawn*, which is concluded with the *Epoch of Reionization*.

By 'Universe', we here mainly refer to its main bright (as opposed to dark) constituent, the gas residing in between the galaxies (Meiksin, 2009). Furthermore, by 'gas' we refer to hydrogen and helium, the main constituents of this intergalactic medium (IGM).

We will first in section 1.1 present the cosmological landscape, in section 1.2 delve deeper into the IGM and what is known about its thermal and ionization history, before turning in section 1.3 to the ionizing and heating sources that have populated it, some since shortly after the beginning of times, in section 1.5 we detail how the radiative properties of these sources can be used theoretically to predict heating, and finally in section 1.5 we will present the main theory behind ionizing numerical radiative transfer which is the method that allows us to answer our two questions. In the subsequent chapters we will provide our answers, and in the concluding chapter, we will at last be ready to present the fascinating story of the Cosmic Dark Ages.

1.1 Cosmos

We will begin this thesis in a placid but controversial astronomical landscape. We have reached the era which is known to be governed by 'precision cosmology'. The cosmological feuds of the 20th century are by and large settled. The proponents of a bewilderingly expanding Universe and as much those that resorted to the comforts of a steady-state, static Universe are all gone. The mystery that the redshifting of photons once was, is now interpreted as a signature of the expansion of space.

With the general theory of relativity, Einstein (1917) paved way for a paradigm shift in the natural sciences, fundamentally changing our notion of time and space. In astronomy, it paved way for new cosmologies. At the time when the precursor to the present-day concordance cosmology had an interpretation where galaxies were point sources creating matter, Hoyle (1948) imagined a continuous and ubiquitous creation of it. A consequence was that it should exist in between galaxies as well. This idea spurred observational searches for an IGM that were bound to be fruitless for another half a century.

However, one of the earliest inquiries into the whereabouts and existence of the IGM was that of Penzias & Wilson (1969) which led to the accidental detection of the blackbody cosmic microwave background radiation (CMB, Penzias & Wilson, 1965). Follow-up observations and theoretical inquiries into its properties have led to believe that the Universe is containing a source of dark energy, conveniently fitting into the framework of Einstein's field equations as a cosmological constant Λ. The Universe is also believed to contain vast amounts of still undetected matter, which could explain the rotational velocity of galaxies (Zwicky, 1933; van de Hulst et al., 1957; Rubin et al., 1980) and that the large scale structure of the Universe appears to be clustered (Einasto et al., 1980), assembling as if a biased tracer of an underlying invisible dark matter field (e.g. Davis et al., 1985; Frenk et al., 1988; Wang et al., 2020). This dark matter is likely to be cold, as warm or hot dark matter would inhibit the formation of observed small-scale structures (e.g. Stücker et al., 2018). These discoveries have led to the present-day cosmology of choice, the ΛCDM-model.

The evolution of the ΛCDM cosmology is expressed in terms of the interpreted expansion, embodied in the dimensionless scale factor a, due to its two main constituents—the cosmological constant Λ and cold dark matter (CDM, subscripted m in equations). The present day value of the scale factor is taken to be 1, and it can be related to a Doppler-like reddening or relativistic time dilatation as $a = (1 + z)^{-1}$ where z is the amount of reddening/dilatation, the redshift, which increases with distance (and light travel time) from us. The prevalence of the constituents at a given point in time can be measured as a ratio $\Omega_x = \rho_x/\rho_c$ of their density ρ_x to the critical density ρ_c, where $x = \{m, \Lambda\}$. The critical density delineates the boundary between a closed ($\rho < \rho_c$), constant ($\rho = \rho_c$) and an open ($\rho > \rho_c$) Universe given a value of the total density ρ, with the value being

$$\rho_c = \frac{3H_0^2}{8\pi G} \tag{1.1}$$

where G is the gravitational constant and H_0 is the present-day interpreted distance-dependent Hubble expansion, in this thesis taken to be 70.1 km s^{-1} Mpc^{-1}. The Hubble

constant above has likely changed with time. The rate the logarithm of the scale factor changes with time—the redshift-dependent Hubble parameter—is

$$H(a) = \frac{d \ln a}{dt} = H_0 \sqrt{\Omega_m a^{-3} + \Omega_\Lambda} \tag{1.2}$$

The Hubble parameter given here is approximately correct for the times explored in this thesis.

We may speculate, perhaps philosophically as to what properties the early Universe should have, and if these properties might be reflected upon the cosmological models. If we instead of dismissing the notions of an Universe which violates energy conservation with its increasing energy, the arrow of time and the unknown nature of the cosmological constant, we could consider these to be smoking guns. The mysterious arrow of time from the second law of thermodynamics, that entropy is increasing, might for example resolve the singularity of big bang. Considering

$$S = -\sum_i p_i \log_2 p_i \tag{1.3}$$

where S is the (information) entropy where we sum over the probabilities $p_i = 1/\Omega_i$ of the configurations Ω_i of the states, we may instantly infer that the beginning of times $S = 0$ corresponds to the Universe being in a single state. It acts as an upper limit on the configuration of the Universe. We could for example interpret it as a manifold M with an associated algebra. It would be simple. All information it can convey is the existence of the Universe so as to not violate $S = 0$. A Universe being in a state with an arrow of time is perhaps also holding a (rather empty, as S must be constant) state in which the arrow of time is oppositely directed—towards lower entropies. Such a reasoning predicts that half of the Universe should be in an inaccessible state, half as so as to maximize the entropy. It would arrive as a coordinate-independent quantity (a constant) in a Lagrangian density describing the Universe, along with the energy densities of inaccessible states other than our own. The further back we go, the more elementary the Universe should be.

However, this picture fits well with the presently accepted model for the initial phase of the Universe. In it, there is no singularity, and the beginning is the '*the extreme opposite of Misner's initial "chaos"*'—according to Starobinsky (1980), the Universe started out in a non-singular way, free of particles and in complete symmetry. It was followed by a rapid inflationary phase (Guth, 1981; Linde, 1982) where 'forces' decoupled, in the process freezing out the elementary particles. Eventually, as these had interacted through the Big Bang Nucleosynthesis to create a handful of heavier particles, the Universe was left in a state of cooling. After a few hundred thousand years, charged particles had combined—first helium, then hydrogen, and light could now stream freely. And finally, the Dark Ages began.

The matter of the cosmological model of Equation 1.2 is predominantly dark, with observable, bright baryonic matter comprising less than ten percent of the mass inferred from observations of the evolution of H. It is this matter that is at focus in this thesis.

1.2 An Intergalactic Medium

The matter that Hoyle (1948) envisioned to exist in between galaxies was thought to likely be hydrogen, which observationally could be revealed either as a deep trough blueward of the HI 21 cm line (Davies & Fennison, 1964) or the HI Lyα line (Gunn & Peterson, 1965) in the spectra of distant, bright radio galaxies (Field, 1962; Penzias & Wilson, 1969; Allen, 1969). However, the initial search for such a medium among the few distant galaxies known at the time was rather fruitless, leading to the conclusion that the IGM must presently be highly ionized and hot.

Another half a century of search (e.g Penzias & Wilson, 1969; Field, 1972; Schneider et al., 1991) has not only ultimately uncovered the IGM's existence (e.g. Fan et al., 2000, 2006; McGreer et al., 2015), but also coincidentally revealed the afterglow from the primordial cosmic fireball (CMB, Penzias & Wilson, 1965). The existence of the IGM is now well-founded by means of various probes such as the disappearance of Lyman α radiation at higher redshifts (Dijkstra, 2014), CMB scattering off electrons freed in the EoR (Planck Collaboration et al., 2018), absorption in the spectra of gamma ray bursts and QSOs (Ciardi & Ferrara, 2005), as well as the intrinsic 21 cm glow of the neutral IGM (Madau et al., 1997). These indicate that the IGM transition from neutral to ionized occurred more than 12 billion years ago. The details of the phase transition, however, remain an outstanding and heavily investigated question. Undoubtedly, something must have driven the phase change by means of providing ionizing and heating radiation that could alter the state of the intergalactic gas. We now turn to some prime candidates for such radiation.

1.3 Sources of Heat and Light

In spite of its name, there were doubtlessly structures forming during the Dark Ages, and some of these produced light. The assembly mechanism of the first bound structures and the time this occurred still remain uncertain. In the young, pristine Universe, conditions were ripe for the rapid formation of very massive objects. In the absence of metals and molecular dust that yet had to be formed, collapsing and heating gas could not cool as effectively as today. However, in the absence of soft UV-radiation that could destroy molecular hydrogen (H_2) or hydrogen deuteride (HD), these molecules provided the one, but inefficient, pathway for gas to condense to form the first generation of stars, PopIII objects (see e.g. the review by Ciardi & Ferrara, 2005, updated 2008-version on arXiv). In the presence of an external, illuminating field that could destroy H_2 and HD, even more massive objects could have formed. These would eventually directly collapse into black holes (Dijkstra, 2014). Their accretion disks would be powerful sources of ionizing radiation, just as the high mass of PopIII stars make them highly luminous.

However, the abundance of these early sources of light is unlikely to have been large enough to have heated and ionized the entirety of the IGM. Instead, we have in this thesis skipped past the first 100 million years in the cosmological evolution to the phase where

the second- and third generation of stars were formed, along with other source types which we are familiar with from the local Universe. In this work, we thus investigate the impact of the following source types.

Stars residing within galaxies. These are thought to be the primary drivers of reionization, as argued by Madau et al. (1999). The recent low Thomson scattering optical depth reported by the Planck Collaboration et al. (2016), combined with constraints from Hubble (Robertson et al., 2015), has rejuvenated the interest in their effect on reionization. Structure formation is also affected by the thermal and ionizing feedback from stars, as discussed by e.g. Couchman & Rees (1986), Cen (1992), and Fukugita et al. (1994). Depending on their initial mass function, mass and metallicity, they will produce copious amounts of ionizing photons, as shown in the review by Ciardi & Ferrara (2005).

Stellar type sources are believed to be the dominant driver of the EoR if a sufficiently large fraction (parametrised as the escape fraction $f_{\rm esc}$) of the ionizing stellar radiation generated in galaxies escapes, but observations indicate that a wide range of escape fractions are possible. It is not clear whether stars could drive reionization if $f_{\rm esc}$ is as low at high-z as it is found to be in many low-z galaxies (see e.g. Vanzella et al., 2016, 2018; Matthee et al., 2018; Naidu et al., 2018; Steidel et al., 2018; Fletcher et al., 2019, for $f_{\rm esc}$ inferences).

Accreting nuclear black holes (or simply black holes, BHs). Hereafter this term will refer to active galactic nuclei (AGN) and the brighter QSOs. Black holes (BHs) have been prime candidates for the ionization of the Universe (e.g. Rees & Setti 1970 and Arons & McCray 1970) ever since the early days of the unavailing search for the intergalactic medium (IGM, e.g. Field 1959). With the detection of 22 faint BHs at $z > 4$, Giallongo et al. (2015) revived the question of their role in this process. It is theoretically possible that BHs could complement low-$f_{\rm esc}$ stellar radiation in reionizing hydrogen in a 'concerted' manner (e.g. Kakiichi et al., 2017). There is thus room for BHs even in the picture of a stellar-dominated EoR. The question remains how large their contribution is.

The most extreme scenario is one where BHs are solely responsible for driving reionization. It was explored by Madau & Haardt (2015), while respecting the Giallongo et al. (2015) constraints on the BHs' overall luminosity. They could positively match the evolution of the volume filling factor of ionized hydrogen, $Q_{\rm HII}$, so as to fit with observations (e.g. Bouwens et al., 2015b). This scenario however fails at reproducing several other observations. BHs alone would yield IGM temperatures and heating that are too high (see e.g. the comparison of BH-only models to the compilation of IGM temperatures of Garaldi et al., 2019), which is followed by too early adiabatic cooling. Furthermore, in the BHs dominated model of Madau & Haardt (2015), HeII reionization would be completed prematurely at $z \sim 4.2$, shortly after HI reionization, which is at odds with the observed extended HeII reionization process (Worseck et al., 2016, 2019). The observational constraints on the ionizing ouptut from high-z

(e.g. Onoue et al., 2017; Parsa et al., 2018; Matsuoka et al., 2018; Kulkarni et al., 2019), as well as theoretical inferences (e.g. Finkelstein et al., 2019), indicate that BHs supply a significant contribution to the ionizing budget, albeit subdominant to that of stars. As discussed by D'Aloisio et al. (2017), BHs can provide an elegant explanation to a flat redshift evolution of the ionizing emissivity, justify the low optical depths in the HeII Lyman α forest, and importantly, explain the origin for the large variations in the opacity of HI Lyman α forest along different sightlines (as investigated by e.g. Chardin et al., 2015).

The interplay between BHs and their host galaxies shapes them both (e.g. Di Matteo et al., 2005). Observations have revealed that massive BHs exist already by $z = 7.5$ (Bañados et al., 2018; Fan et al., 2019), and simulations do not rule this out as unfeasible (e.g. Feng et al., 2015; Di Matteo et al., 2017). The growth of BHs can be captured well by simulations (e.g. Sijacki et al., 2015; DeGraf et al., 2012; Weinberger et al., 2018; Huang et al., 2018), however, the question of their formation remains still open (e.g.; Volonteri & Bellovary, 2012; Regan & Haehnelt, 2009, for recent reviews). A common numerical approach in large cosmological volume simulations (Di Matteo et al., 2012; Khandai et al., 2015; Sijacki et al., 2015; Crain et al., 2015; Weinberger et al., 2018), is to seed galaxies above a mass threshold with a BH of mass close to the mass resolution (typically BH seeds of $10^{4-5}\mathrm{M}_{\odot}$ within halos of $10^{10-11}\mathrm{M}_{\odot}$). This approach leads to a population of BHs at $z = 0$ that matches observations (e.g. Kormendy & Ho, 2013, for a recent review), however it does not shed light on the abundance and properties of faint/small mass BHs at higher z, a population which can be important during the initial stages of the EoR.

Galactic X-ray binary systems (XRBs). These comprise a neutron star or a black hole devouring a companion star. Among such systems, the majority of the ionizing luminosity at high-z originates from massive (HMXBs) rather than low-mass (LMXBs) binary systems (Mirabel et al., 2011; Fragos et al., 2013b,a; Madau & Fragos, 2017; Sazonov & Khabibullin, 2017). Mineo et al. (2012a,b) found the spectra of XRBs to be too hard to account for the soft X-ray flux of galaxies, while they become dominant at higher energies.

While XRBs dominate the X-ray output of gas-poor galaxies (Fabbiano, 2006; Mineo et al., 2012a), their hard spectra are not necessarily primarily accompanied by IGM ionization, but rather by heating (Fialkov et al., 2014b).

Thermal bremsstrahlung from the interstellar medium (ISM). The heating mechanism of the diffuse gas has been a topic under investigation for decades (see e.g. the review by Fabbiano 1989). Interactions between supernova driven galactic superwinds and clouds (e.g. Chevalier & Clegg, 1985) is the preferred explanation (Pacucci et al., 2014). Shocked gas in the galactic halo and disk (e.g. Suchkov et al., 1994), as well as hot galactic winds (e.g. Strickland & Stevens, 2000) are also processes that could yield predominantly soft X-rays.

The softer spectrum of the ISM potentially likens its reionization signatures to those of cosmic rays (e.g. Sazonov & Sunyaev, 2015; Leite et al., 2017; Jana & Nath, 2018) and AGB stars (Vasiliev et al., 2018).

Other candidates. These will not be examined further due to their secondary role. Such sources include e.g. low energy cosmic rays (Ginzburg & Ozernoi, 1965; Nath & Biermann, 1993; Sazonov & Sunyaev, 2015; Leite et al., 2017), self-annihilation or decay of dark matter (e.g. Liu et al. 2016) and plasma beam instabilities in TeV blazars (e.g. Chang et al. 2012; Puchwein et al. 2012).

1.4 Ionization and Heating

A source of radiation may also be a source of heat if the radiation loses energy to the medium it interacts with. We are for example familiar with heating due to molecular excitement of water in microwave ovens or from infrared radiation. The primary means of heating the intergalactic medium is through its ionization. The ionization rate Γ_X of an atom/ion X is dependent on the irradiating spectrum, expressed here through the specific intensity J_ν in ergs s^{-1} Hz^{-1} cm^{-2} sr^{-1}, an angle averaged integral of the intensity I_ν, $J_\nu = 1/(4\pi) \int I_\nu \, \mathrm{d}\Omega$,

$$\Gamma_X = \int_{\nu_X}^{\infty} \frac{4\pi J_\nu}{h_\mathrm{P} \nu} \sigma_{X,\nu} \, \mathrm{d}\nu \tag{1.4}$$

where h_P is Planck's constant, ν_X the ionization threshold frequency and $\sigma_{X,\nu}$ the frequency dependent ionizing cross section. The integral is from the ionization threshold ν_X to infinity, as lower energy radiation is unable to ionize the ion. The heating $\mathcal{H}$ is the excess energy leftover from ionizations, expressed as the photoheating rate G in ergs s^{-1},

$$\mathcal{H} \approx G = \sum_X p_X n(X) \int_{\nu_X}^{\infty} \frac{4\pi J_\nu}{h_\mathrm{P} \nu} h \left(\nu - \nu_X\right) \sigma_{X,\nu} \, \mathrm{d}\nu \tag{1.5}$$

where we sum over the different species X whose ionization probability is p_X and number density is $n(X)$ in cm^{-3}. The temperature increase ΔT_X of a species X following the passage of an ionization front was estimated by Abel & Haehnelt (1999) to be roughly

$$k_\mathrm{B} \Delta T_X \approx \frac{G_X}{\Gamma_X} \approx \frac{h_\mathrm{P} \nu_X}{\alpha_J + \alpha_\tau} \tag{1.6}$$

where k_B is Boltzmann's constant and α_J and α_τ being two spectral indices. We have α_J which depends on the power-law slope of the incident spectrum J_ν at the ionization threshold ν_X and $\alpha_\tau = 2$ in an optically thin medium (with an optical depth $\tau < 1$, where $\tau = N_X \sigma_{X,\nu}$ is a function of the column density N_X in cm^{-2} and the cross section $\sigma_{X,\nu}$) and $\alpha_\tau = -1$ in an optically thick medium ($\tau > 1$). This relation does however not account for the pronounced increase in number density with a. This effect attenuates the temperature

significantly. We have indications that a correction of Equation 1.6 into (Fiaschi et al., in prep.)

$$k_{\rm B}\Delta T_X \approx \frac{1}{(1+z)^3}\frac{h_{\rm P}\nu_X}{\alpha_J+\alpha_\tau} \tag{1.7}$$

yield predictions more in line with simulations. Furthermore, as we will return to, we have found that stellar-type spectra are much harder near the hydrogen ionization threshold than what has been previously assumed in literature, where $\alpha_J \sim 3$.

Gas will also undergo cooling. Such processes include situations in which the electron recombines cascading through the energy levels of an ion (case B recombination) and low energy photons are released. A direct recombination of an electron to the ground state of an ion (case A) will emit a photon sufficiently energetic to cause a new ionization event. In the presence of positively charged particles that move relative to electrons, the latter, which are lighter, will experience a trajectory-changing acceleration due to the emergent magnetic field from this relative motion. This is a free-free process between free particles, Bremsstrahlung, in which the particle deceleration is accompanied by a radiative loss with a scale-free power-law like spectrum. A free electron may also cool through interactions with neutral particles. In the low energy regime, the electron may simply collisionally excite the ion. In the high energy regime, it may however ionize it, causing a secondary ionization event (Shull & van Steenberg, 1985). The free electrons may also deposit energy onto a background radiation field through inverse Compton scattering, such as in the Sunyaev-Zeldovich effect (Sunyaev & Zeldovich, 1972).

It should be noted that Equation 1.6 is an approximation. The correct change in temperature can be expressed as

$$\frac{dT}{dt} = \frac{2}{3k_{\rm B}T}\left[k_{\rm B}T\frac{dn}{dt} + \mathcal{H} - \Lambda\right] \tag{1.8}$$

following Maselli et al. (2003). We have the differential change in gas temperature T with time t in seconds which increases if the number density n increases, occurring when ionizations or shocks liberate particles, and is countered by the cosmological decrease in n with a^{-3}. The second term corresponds to the aforementioned heating in Equation 1.5. The third term, Λ in ergs s^{-1}, is the cooling rate and embeds the cooling processes mentioned above. Expressions for these rates can be found in e.g. Cen (1992).

1.5 Simulating Heating and Ionization

To answer the overarching questions of this thesis, we have to resort to numerical simulations that can mimic the relevant processes contributing to the cosmic intergalactic heating and reionization. Ideally, the radiative transfer process is coupled to the simulation of structure formations as the heating and reionization process influences the former (see e.g. Katz et al., 2020). This in turn influences the abundance and properties of the sources that emit ionizing radiation. However, this coupling is computationally expensive

and prohibits us from exploring volumes large enough to be representative of cosmos (Iliev et al., 2014) and with sufficient radiative transfer accuracy to capture the physics and influencers of the non-linear process that heating and ionization is.

The central equation that is sought solved is the radiative transfer equation, which in its simplest form can be written as

$$\frac{\mathrm{d}I(\nu)}{\mathrm{d}\tau(\nu)} = S(\nu) - I(\nu) \tag{1.9}$$

in which we have a change in the monochromatic specific intensity $I(\nu)$ with a change in optical depth $\tau(\nu)$ due to radiative contributions in the source function $S(\nu)$ in ergs s^{-1} Hz^{-1} cm^{-2} sr^{-2},

$$S(\nu, s) = \frac{j(\nu, s)}{\alpha(\nu, s)} + \iint \frac{\alpha(\nu', s)}{\alpha(\nu, s)} I(\nu', s') R(\nu, \nu', s, s') \, \mathrm{d}^3 s' \, \mathrm{d}\nu' \tag{1.10}$$

where the first term on the right hand side is the ratio between the emission coefficient $j(\nu, s)$ in ergs s^{-1} cm^{-3} Hz^{-1} sr^{-1} and the monochromatic linear extinction coefficient $\alpha(\nu, s)$ in cm^{-1}. The emission coefficient relates to the volume emissivity ε in ergs s^{-1} cm^{-3} Hz^{-1} as $\varepsilon(\nu, s) = 4\pi j(\nu, s)$. The second term contains the redistribution function $R(\nu, \nu', s, s')$ in Hz^{-1} cm^{-3} which gives the probability of redistribution from one frequency ν' to ν from the direction s' into s.

The source terms embed processes which represent either locally produced radiation (emissivity) or scattered radiation (redistribution). The sink term, on the other hand, is due to radiative losses when the optical depth increases. This interpretation also sheds light on the source term—radiation only appears to emanate from a medium that is not transparent.

We also need to account for redshifting. Doppler shifting of radiation interacting with a moving medium is a striking example of how radiation can be redistributed. However, cosmological redshifting is not due to interactions, but rather to the relativity of time—from our vantage point it appears to pass slower at higher redshifts. To account for this, we may rewrite the change in the optical depth as

$$\mathrm{d}\tau = \alpha \, \mathrm{d}s = \alpha c \, \mathrm{d}t \tag{1.11}$$

and as this is a total derivative, we have to take greater care with the dependencies of the intensity,

$$I = I(\nu, \hat{\mathbf{n}}(\mathbf{x})) \tag{1.12}$$

where we have specified the line element $\mathrm{d}s$ to be the normal vector $\mathrm{d}\hat{\mathbf{n}}$ at the point $\mathbf{x}$. This leaves the following expression for the total (optical) derivative,

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{1}{\alpha c}\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\dot{a}}{\alpha c}\frac{\mathrm{d}}{\mathrm{d}a} = \frac{\dot{a}}{\alpha c}\left[\frac{\partial}{\partial a} + \frac{\partial \nu}{\partial a}\frac{\partial}{\partial \nu} + \sum_i \left(\frac{\partial \hat{n}_i}{\partial a}\frac{\partial}{\partial \hat{n}_i} + \frac{\partial x_i}{\partial a}\frac{\partial}{\partial x_i}\right)\right] \tag{1.13}$$

where we made use of the relation between the scale factor and its derivative, $\dot{a} = \mathrm{d}a/\mathrm{d}t$. We may make use of $\nu \propto a^{-1}$, and hence $\partial\nu/\partial a = -\nu a^{-1}$; $\hat{n} \propto a$ and hence $\partial\hat{n}_i/\partial a = 1$; and finally that $\partial r_i/\partial a = \dot{a}^{-1}\partial r_i/\partial t$. This leaves

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \frac{1}{\alpha c}\frac{\partial}{\partial t} + \frac{1}{\alpha c}\frac{\dot{a}}{a}\left[-\nu\frac{\partial}{\partial\nu} + 3\right] + \sum_i \frac{1}{\alpha c}\frac{\partial x_i}{\partial t}\frac{\partial}{\partial x_i} \tag{1.14}$$

where we have the temporal derivative followed by a bracketed term corresponding to frequency redshifting and Hubble flow, and finally a familiar divergence term, where we may use that the divergence term in vector units are $(\alpha c)^{-1} c\hat{\mathbf{n}}\cdot\boldsymbol{\nabla}$. This brings our equation in line with e.g. Meiksin (2009). Our cosmological radiative transfer equation is now found by simply combining Equation 1.14 with Equation 1.9,

$$\frac{1}{\alpha(\nu)c}\frac{\partial I(\nu,\hat{\mathbf{n}}(\mathbf{x}))}{\partial t} + \frac{1}{\alpha(\nu)c}\frac{\dot{a}}{a}\left[3I(\nu,\hat{\mathbf{n}}(\mathbf{x})) - \nu\frac{\partial I(\nu,\hat{\mathbf{n}}(\mathbf{x}))}{\partial\nu}\right] + \frac{1}{\alpha(\nu)}\hat{\mathbf{n}}\cdot\boldsymbol{\nabla}I(\nu,\hat{\mathbf{n}}(\mathbf{x})) = S(\nu,\hat{\mathbf{n}}(\mathbf{x})) - I(\nu,\hat{\mathbf{n}}(\mathbf{x})). \tag{1.15}$$

In the context of hydrogen and helium reionization, we can safely ignore scatterings and kinematic Doppler shifts of our radiation. The more readily ionized hydrogen is more abundant in the IGM than than helium which has a higher ionization threshold, and hydrogen ionizing radiation is more likely to be lost to ionizations than for it to experience helium scattering. If we furthermore only consider the ionizing radiation to originate in 'point-like' sources such as galaxies (on intergalactic scales), and disregard diffuse ionizing recombination radiation, we may set $j(\nu, s) = \varepsilon(\nu, s')\delta(s - s')/(4\pi)$ where we now have a (point) source at s'. For paths along s not containing point sources, Equation 1.9 simplifies to

$$\frac{\mathrm{d}I}{I} = -\mathrm{d}\tau \tag{1.16}$$

with the solution

$$I(\nu,\tau) = I(\nu,0)\exp\left[-\int_0^{\tau(\nu,s)}\mathrm{d}\tau'\right] = I(\nu,0)\exp\left[-\int_0^{s}\alpha(\nu,s')\,\mathrm{d}s'\right] \tag{1.17}$$

where we have refrained from assuming that the extinction coefficient is constant along the path we are integrating. The meaning of the equation is however clear—as we progress from a point $s' = 0$ to s the radiation is attenuated naturally according to the amount of extinction it experiences along the path.

Translating this attenuation into ionization and heating can be done by considering a radiation fluid which propagates diffusively by integrating Equation 1.15 numerically (e.g. Gnedin & Abel, 2001; Cen, 2002; Hayes & Norman, 2003; Kannan et al., 2019), or in a Monte Carlo fashion by discretizing the radiation field into individual rays originating either at the sources or the cells which deposit their energy as they propagate and encounter neutral gas (e.g. Norman et al., 1998; Ciardi et al., 2001; Mellema et al., 2006a).

In this thesis we use the multifrequency 3D Monte Carlo radiative transfer code `CRASH`, the *Cosmological RAdiative transfer Scheme for Hydrodynamics* (Ciardi et al., 2001; Maselli et al., 2003, 2009; Graziani et al., 2013; Hariharan et al., 2017; Graziani et al., 2018; Glatzle et al., 2019). Central in this approach is our ability to discretize the radiation field in time, direction and frequency. The spectral shape of the sources is discretized into 82 frequency bins, where 72 are in the UV regime below $h_{\rm P}\nu = 200$ eV, and the rest are in the soft X-ray regime (0.2–2 keV). We choose to space the bins densely around the ionization thresholds of hydrogen (13.6 eV) and helium (24.6 eV and 54.4 eV).

The radiation emitted by each source i at a redshift z is dicretized into N_γ photon packets (see Appendix A.1 for a convergence analysis motivating this choice). Each packet, $N_{{\rm ph},i}(\nu, z)$ (in units phots Hz^{-1}) holds the total number of photons emitted by source i during the timestep $\Delta t_{\rm em}(z)$ (in s) in the different frequency bins:

$$N_{{\rm ph},i}(\nu, z) = \frac{S_i(\nu, z)}{L_i(z)} f_{\rm esc}(\nu)\varepsilon_i(z)\Delta t_{\rm em}(z), \tag{1.18}$$

where $f_{\rm esc}(\nu)$ is the frequency-dependent escape fraction, $\varepsilon_i(z)$ is the rate of ionizing photons the source emits (in phots s^{-1}), $S_i(\nu, z)/L_i(z) = \hat{S}_i(\nu, z)$ is the normalized spectral shape (in Hz^{-1}), where $S_i(\nu, z)$ defines the Spectral Energy Distribution (SED; in units of ergs Hz^{-1} s^{-1}) and $L_i(z)$ the luminosity (in units of ergs s^{-1}). The emissivity, the emission rate of ionizing photons, can be written as:

$$\varepsilon_i(z) = \int_{13.6{\rm eV}}^{2{\rm keV}} \frac{S_i(\nu, z)}{h_{\rm P}^2 \nu}\, {\rm d}(h_{\rm P}\nu). \tag{1.19}$$

We adopt a single escape fraction $f_{\rm esc}$[1] of UV photons for all sources at a given redshift, except for BHs, for which photons at all frequencies are assumed to escape. Although the adoption of a single value for the escape fraction is an over simplification as in reality it depends e.g. on the mass of dust and gas and redshift of the host galaxy, its density distribution, the mass and location of the stellar sources (see e.g. Ciardi & Ferrara 2005 and its updated version on arXiv), $f_{\rm esc}$ is to a large degree an unconstrained parameter at low (see e.g. Matthee et al. 2017 and Vanzella et al. 2016) as well as high redshift, where physical conditions could have promoted large values of $f_{\rm esc}$ (see e.g. Kitayama et al. 2004; Yoshida et al. 2007; Safarzadeh & Scannapieco 2016 or the recent $z = 4$ observations by Vanzella et al. 2018).

We have now presented the backdrop for the remainder of this thesis. We have explored cosmos, the theoretical origins of the IGM, some radiative sources that could have ionized and heated the Universe, and we have presented analytic and numerical ways of investigating the two fundamental questions we started out by asking. In the following, we will first present simulations of heating and ionization during the early Cosmic Dawn in chapter 2, before turning to presenting a full reionization history in chapter 3, and finally asking the question whether our BH model is complete, or whether we are missing out on

[1] `CRASH` allows us to specify the escape fraction in different frequency bands individually for each source.

a large fraction of faint, light nuclear BHs in chapter 4. We conclude with the story of the Dark Ages in chapter 5.

Chapter 2

The Epoch of Cosmic Heating

2.1 Introduction

As put forth in chapter 1, the early temperature of the IGM is a a major uncertainty in the evolution of the Universe. The heating efficacy of the sources depends on their spectral softness, a central finding by e.g. Ciardi et al. (2012) and Fialkov et al. (2014b). This is something we also expect from a theoretical point of view, as discussed in chapter 1. However, a complicating factor is the fact that the different source types have different spectral properties, and that all are likely to be present during the Dark Ages to produce radiation that may affect the IGM's thermal history. This 'concerted picture' of the sources' potential relevance was discussed by Kakiichi et al. (2017). The main contribution of this chapter to the canon of IGM literature is our examination of the relative role the various sources may play in the preheating of the IGM during the Dark Ages. The work presented here is published in Eide et al. (2018).

Early cosmological 21 cm surveys have begun to report that heating may precede full ionization, resulting in a separate *Epoch of Heating* (EoH) preceding the *Epoch of Reionization* (EoR). The recent constraint on the global signal with SARAS 2 (Singh et al., 2017) disfavors late heating, disentangling the EoH from the EoR (Fialkov et al., 2014b). These observations also align themselves with other surveys, as the early results of Bowman et al. (2008), that were followed by upper limits on the scale and magnitude of brightness temperature fluctuations (e.g. from GMRT, Paciga et al. 2011; MWA, Dillon et al. 2014; Ewall-Wice et al. 2016; PAPER, Parsons et al. 2014; Ali et al. 2015; and LOFAR, Patil et al. 2017). However, these surveys are impeded at very high redshift by the ionospheric contamination from Earth, which is nearly fully opaque to redshifted 21 cm signal. To mitigate for this, a seminal Chinese-Dutch mission launched in 2018 plans to obtain the global 21 cm signal from the far side of the Moon[1].

Besides 21 cm tomography, which is still in its infancy after its revival as a probe by Madau et al. (1997), there are other observational means for examining the ionization and thermal state of the IGM. The spectra of bright, single sources (as quasars or gamma ray

[1]https://www.isispace.nl/dutch-radio-antenna-depart-moon-chinese-mission/, accessed 13 July 2017.

bursts, see e.g. review by Ciardi & Ferrara, 2005) reveal the properties of the intervening IGM along the line-of-sight. The increasing abundance of neutral HI at earlier z has been indicated from large-scale surveys, as well (e.g. Matthee et al. 2015; Zheng et al. 2017; Ouchi et al. 2018 or see review by Dijkstra 2014). Hydrogen ionization is accompanied by production of free electrons. Satellite-based experiments (Komatsu et al., 2011; Planck Collaboration et al., 2016) have provided tight constraints on the scattering of the cosmic microwave background (CMB) off these electrons, as the CMB becomes dampened and linearly polarized. These observations are however in slight tension, as the continuing lowering of the Thomson scattering optical depth requires a later completion of cosmic reionization.

There are different approaches for simulating the physics of the Cosmic Dawn. Ideally, structure formation can be coupled to radiative transfer (e.g. Gnedin 2014; So et al. 2014; O'Shea et al. 2015; Ocvirk et al. 2016; Pawlik et al. 2017; Semelin et al. 2017). However, this is computationally expensive and requires simplifications. For example helium physics is often not treated, or the number of frequency bins used may be limited to a few, albeit the cross section of hydrogen varies approximately as ν^{-3}. The spatial extent is also often well below the required $\sim 100h^{-1}$ cMpc needed to provide a consistent reionization history (Iliev et al., 2014). Semi-numerical codes as 21CMFAST (Mesinger et al., 2011) or SimFast21 (Santos et al., 2010; Hassan et al., 2016) can be applied on large scales using an excursion-set formalism (Furlanetto et al., 2004). This allows for fast parameter-space exploration, but may lack the treatment of some important physical processes, such as helium reionization, temperature evolution, partial ionization and X-ray implementation. Nevertheless, this approach has proved to be rewarding in studies of possible consequences of radiative feedback on structure formation and its 21 cm signatures (see e.g. Fialkov et al., 2014a). Radiative post-processing of hydrodynamical or N-body simulations is also possible (e.g. Baek et al., 2010; Ciardi et al., 2012; Graziani et al., 2015; Ross et al., 2017) and allows to dedicate computational power to e.g. the treatment of heating, multifrequency radiative transfer (RT), as well as the effects of helium.

Here we examine the effects of *stars, BHs, XRBs* and the *ISM* on both the thermal and ionization states of the IGM during the EoH. Our approach is the following: we rely on the hydrodynamical cosmological structure formation simulation MassiveBlack-II (Khandai et al., 2015, MBII, described in §2.2.1), which includes baryonic physics and feedback processes, to provide us with the physical environment of the IGM (temperature and gas density), as well as the location and properties of the sources. We assign spectra and ionizing luminosity to star forming particles based on their star-formation rates, masses, ages and metallicities, and to black hole particles based on their accretion rates. We then perform radiative post-processing of the MBII simulations with the cosmological radiative transfer code `CRASH` (see §2.2.2) which gives us the evolution of the thermal and ionization state of the IGM. The properties of the ionizing sources are based on empirical relations introduced in §2.2.3. We present our findings in §2.3 and relate them to comparable studies also summarizing our conclusions in §2.4 and §2.5.

2.2 Methodology

Here, we describe how we combine the outputs of the cosmological hydrodynamical simulation MBII (§2.2.1) with population synthesis modeling of ionizing sources (§2.2.3), and finally perform multifrequency radiative transfer in the same cosmological volume with `CRASH` (§2.2.2).

2.2.1 Cosmological hydrodynamic simulation

MBII (Khandai et al. 2015, K15) is a high resolution cosmological SPH simulation tracking stellar populations, galaxies, accreting and dormant black holes, as well as their properties (as position, age, metallicity, mass, accretion rate, star formation rate). The simulation has been run using `P-GADGET`, a newer version of `GADGET-3` (see Springel, 2005, for an earlier version). It accounts for baryonic physics and feedback effects of the sources on their environment following earlier works in its approach to feedback (Di Matteo et al., 2008; Croft et al., 2009; Degraf et al., 2010; Di Matteo et al., 2012), sub-grid treatment of star formation (Springel & Hernquist, 2003), and seeding and evolution of black holes (Di Matteo et al., 2005; Springel et al., 2005). The simulation has a box length of $100h^{-1}$ cMpc and is performed in the WMAP7 ΛCDM cosmology (Komatsu et al., 2011)[2], using 2×1792^3 particles of mass $m_{\rm DM} = 1.1 \times 10^7 h^{-1}{\rm M}_\odot$ and $m_{\rm gas} = 2.2 \times 10^6 h^{-1}{\rm M}_\odot$ for dark matter and gas, respectively. The adopted gravitational softening length is $1.85h^{-1}$ ckpc. As a reference, we have a total of 28 and 169,520 haloes hosting at least a star particle at $z = 18$ and 10, respectively, where the lowest dark matter halo masses are $2\times10^8 h^{-1}{\rm M}_\odot$ and $9\times10^7 h^{-1}{\rm M}_\odot$. The highest dark matter halo masses are $2\times10^9 h^{-1}{\rm M}_\odot$ and $1\times10^{11} h^{-1}{\rm M}_\odot$ at $z = 18$ and 10, respectively. The BHs form from seeds of $5 \times 10^5 h^{-1}{\rm M}_\odot$, growing with accretion rates in the range $(10^6$–$10^7)h^{-1}{\rm M}_\odot{\rm Gyr}^{-1}$ to a maximum mass of $1.4\times10^6 h^{-1}{\rm M}_\odot$ at $z = 10$. The first seed black hole is found at $z = 13$, while we have seventeen BHs at $z = 10$. We refer the reader to K15 for more details on the simulation.

We employ six snapshots from MBII, covering the evolution between redshifts $z = 18$ and $z = 10$, each describing the instantaneous state of the simulation. The particle distribution of each snapshot is mapped onto a Cartesian grid of N_c^3 cells to create maps of gas number density and temperature, as well as location and properties of the ionizing sources (see §2.2.3 for a detailed description of how we convert MBII data to ionizing sources). The reference value is $N_c = 256$, corresponding to a spatial resolution of $391h^{-1}$ ckpc. After gridding, the 45 (292,685) star particles present at $z = 18$ (10) are reduced to 26 (56,702). These are effectively our sources. When more than one particle ends up in a cell, their properties are summed up (see following section). The gas density is converted to hydrogen and helium number densities by assuming a number fraction $X = 0.92$ and $Y = 0.08$, respectively, and no metals.

[2] $\sigma_8 = 0.816$, $n_s = 0.968$, $\Omega_\Lambda = 0.725$, $\Omega_{\rm m} = 0.275$, $\Omega_{\rm b} = 0.046$, $h = 0.701$

2.2.2 Cosmological radiative transfer simulation

As detailed in section 1.5, a precise numerical implementation of the radiative transfer (RT) process is required to catch the various physical processes involved when the radiation of our sources heat and ionize the IGM. We refer the reader to chapter 1 for further details of our implementation. In summary, the RT of ionizing photons is performed by post-processing the outputs of our adopted cosmological environment MBII with the multifrequency Monte Carlo ray-tracing code `CRASH` (Ciardi et al., 2001; Maselli et al., 2003, 2009; Graziani et al., 2013), which calculates the ionization state of hydrogen and helium, as well as the gas temperature in each grid cell traversed by photons emitted by a radiation source. The version of `CRASH` employed here features a self-consistent treatment of UV and soft X-ray photons, in which X-ray ionization and heating as well as detailed secondary electron physics are included (Graziani et al., 2018).

In the following section we will describe how the spectrum, luminosity and emissivity of the sources are modeled.

2.2.3 Sources of ionizing radiation

In our simulations we consider four types of ionizing sources:

(i) regular stars, hereafter abbreviated only as *stars*,

(ii) neutron star/black hole X-ray binaries, hereafter *XRBs*,

(iii) thermal bremsstrahlung from supernova-heated ISM, hereafter *ISM*.

(iv) accreting nuclear black holes, hereafter *BHs*.

For each source i of any of the types s we model its SED, S_i^s, and luminosity, L_i^s, as detailed in the following subsections. When several sources of different types are present in the same cell, we sum up their contributions (except for BHs, which are treated separately) to obtain $S_i = \sum_s S_i^s$ and $L_i = \sum_s L_i^s$. Using Eqs. 1.19 and 1.18, we can then evaluate the photon content of each packet emitted.

In Fig. 2.1 we plot the evolution of the median, the 25th to 75th percentiles and the minimum/maximum values of the emissivity for the various source types, as evaluated from Eq. 1.19. In the lower panel, we plot the comoving volume averaged emissivities. These reflect the abundance of the sources, unlike the individual emissivites plotted in the upper panel, showing that stars provide the bulk of the ionizing photons throughout the redshift range considered here. Even though individual BHs have higher emissivities than the majority of the stars in galaxies, they are much fewer in number, and are therefore not dominating the overall emissivity budget. The median values of the emissivities for the different source types remain also fairly constant. This applies to the maximum values too: the brightest stellar-type sources at $z = 14$ are less than an order of magnitude brighter at $z = 10$. The brightest black holes, on the other hand, do not have higher emissivities compared to the brightest galaxies, which can emit an order of magnitude more ionizing

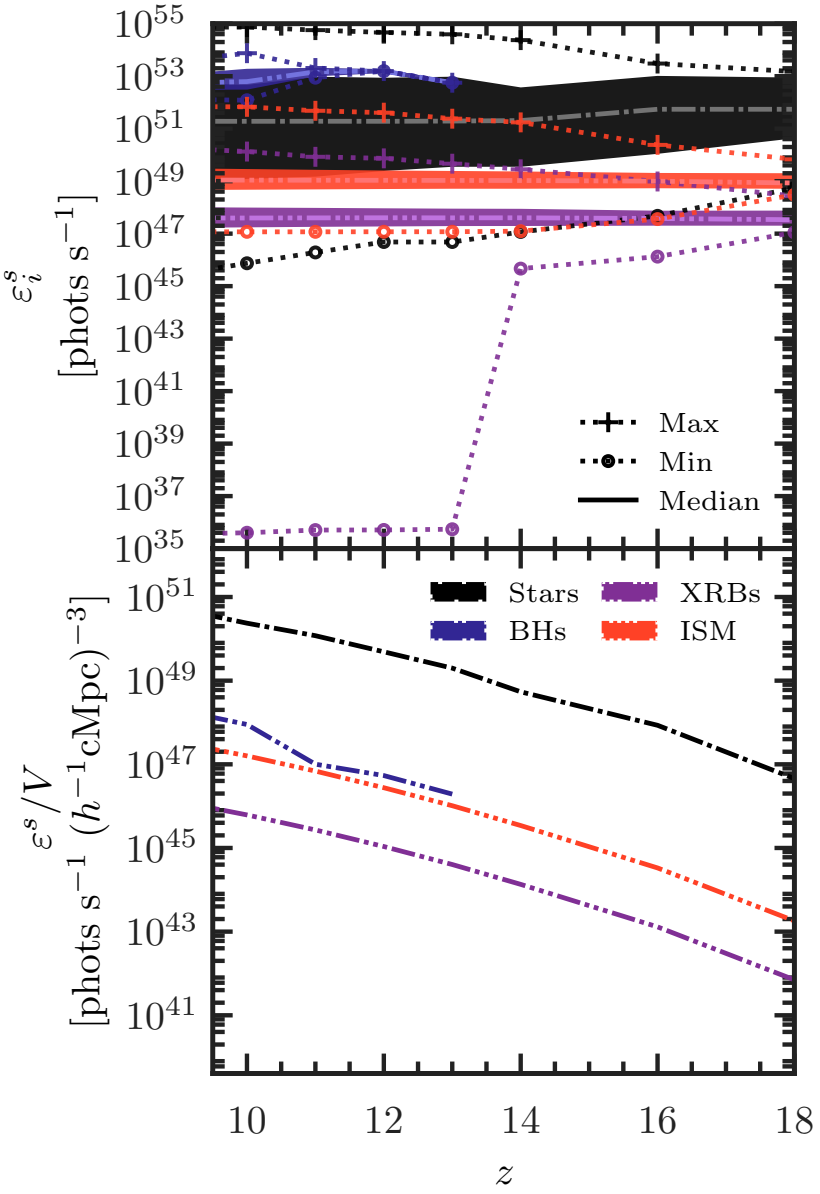

Figure 2.1: *Upper panel*: Redshift evolution of the emissivity ε_i^s of individual sources (upper panel), grouped per source type s. Stars are plotted in black (single dot), BHs in blue (two dots), XRBs in purple (three dot) and the ISM in red (four dots). The solid, central lines are the median values, enclosed in shaded regions that show the 25th to 75th percentile of the values. The span between the minimum and maximum emissivities are given from the dashed lines, with "+" and "o" denoting upper and lower limits, respectively. Note that these values are intrinsic and have not been scaled by an escape fraction. *Lower panel*: Redshift evolution of the volume averaged emissivity ε^s/V per source type s. Here, the escape fractions are accounted for.

photons. The slow evolution in emissivities is related to the physical properties (e.g. stellar masses, star formation rates and metallicities) that goes into determining the SEDs.

We note that the median stellar emissivity decreases with decreasing redshift. Although the median stellar mass increases, suggesting higher luminosities, this is counteracted by an increment of both the median stellar mass weighted metallicity and age of the stars. Nevertheless, stars are the dominant producer of ionizing photons at all redshifts. The evolution of the emissivities of the XRBs and the ISM is dictated by the slow evolution of the star formation rates. The sudden drop in the lower limit of the emissivity of XRBs is due to galaxies that only hosts the fainter LMXBs. The first black hole arises at $z = 13$, and not until $z \leq 11$ do we have more than one present in our volume.

While `CRASH` can handle a different spectrum for each single source, for the sake of simplicity we adopt at each redshift z an average spectral shape for all sources but BHs. More specifically, for each source i we evaluate $S_i = S_i^{\rm stars} + S_i^{\rm XRB} + S_i^{\rm ISM}$. We then use the average $\bar{S} = \langle S_i \rangle$ in Eq. 1.18 rather than S_i. Whenever a BH is present in a cell, this is added as a separate source having the same spatial coordinates. Its spectrum is similarly calculated as $\bar{S}^{\rm BH} = \langle S_i^{\rm BH} \rangle$.

In Fig. 2.2 we plot $\bar{S}$ at different redshifts z. We see that stars dominate at energies $h_{\rm P}\nu \lesssim 60$ eV, while the ISM contribution is relevant above the HeII ionization threshold, i.e. into hard UV and the soft X-rays. The XRBs provide the harder X-rays. The weak redshift evolution is due to a combination of the averaging effect effectively preferring brighter sources, and the mentioned slow evolution in the underlying physical properties determining the spectra.

In the following sections we will describe in more detail how we evaluate the luminosity L_i^s and normalized spectral shape $\hat{S}_i^s$ for the various source types.

Stars

We model the ionizing radiation from stars by using stellar particles identified in MBII. From the age, metallicity and mass of each stellar particle p, we obtain $\hat{S}_p^{\rm stars}$ and $L_p^{\rm stars}$ using the stellar population synthesis code BPASS (Eldridge & Stanway, 2012). We adopt the instantaneous starburst prescription of star formation and the evolution model that does not account for interactions in binary systems[3].

As a single halo i may comprise several particles p, in this case we sum up the contributions from all the particles to obtain $L_i^{\rm stars} = \sum_{p\in i} L_p^{\rm stars}$ and $\hat{S}_i^{\rm stars} = \sum_{p\in i} \hat{S}_p^{\rm stars}$.

X-ray binaries

To account for ionizing radiation coming from X-ray binary systems, we combine the galactic properties provided by MBII with scaling relations from the Fragos et al. (2013b,a) XRB population synthesis model, recently updated by Madau & Fragos (2017). We can thus capture both the metallicity evolution of the high-mass XRBs (HMXBs) and the age

[3]The effect of not including the evolution of binary systems is a reduction in ionizing flux.

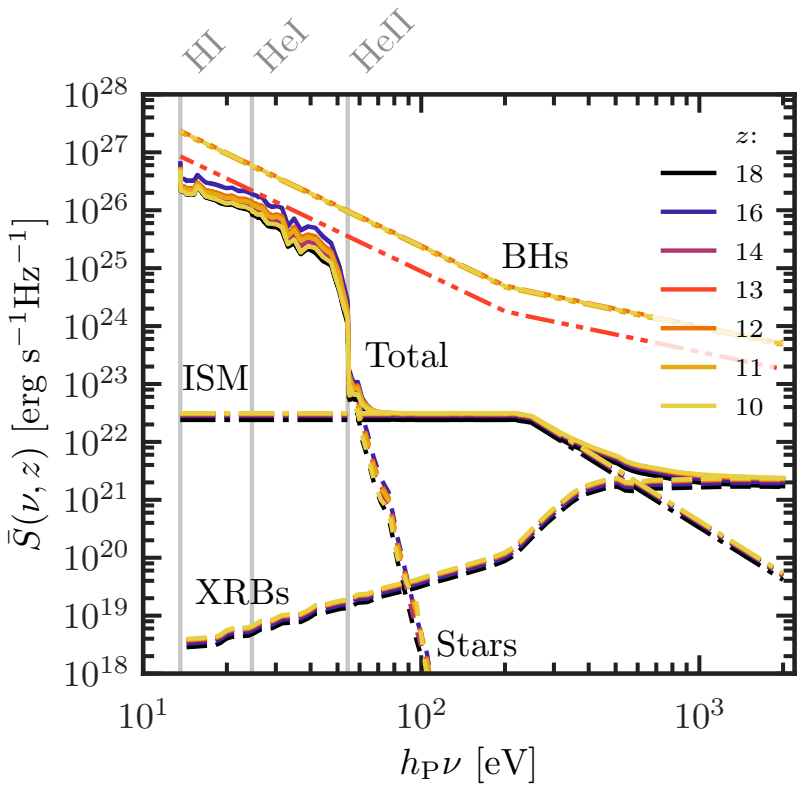

Figure 2.2: Globally averaged spectral energy distributions (SEDs) $\bar{S}(\nu, z)$ of the different source types: stars (short dashed lines), XRBs (long dashed), ISM (short-long dashed), total galactic (solid) and black holes (long/double short dashed). The faint, vertical gray lines indicate the ionization thresholds for hydrogen (13.6 eV), neutral helium (24.6 eV) and singly ionized helium (54.4 eV). Different colors indicate different redshifts.

and stellar mass dependence of the low-mass XRBs (LMXBs), as well as the spectral shape evolution in redshift.

For each stellar particle p, we obtain its spectral shape $\hat{S}_p^{\mathrm{XRB}}$ from Fragos et al. (2013b,a) and its luminosity L_p^{XRB} as:

$$L_p^{\mathrm{XRB}} = L_p^{\mathrm{HMXB}} + L_p^{\mathrm{LMXB}}. \tag{2.1}$$

The contribution from HMXBs can be found using Eq. (3) in Fragos et al. (2013b) with updated coefficients β_{0-4} from Madau & Fragos (2017):

$$\begin{aligned}\log\left(L_p^{\mathrm{HMXB}}/\mathrm{SFR}_i\right) &= \beta_0 + \beta_1 Z_p + \beta_2 Z_p^2 + \beta_3 Z_p^3 + \beta_4 Z_p^4 \\ &\times \left(\mathrm{erg\,s^{-1}\,M_\odot^{-1}\,yr}\right), \quad \text{for } Z \in [0, 0.025]\end{aligned} \tag{2.2}$$

where SFR_i (in $M_\odot\ \mathrm{yr}^{-1}$) is the star formation rate of the halo i hosting particle p, and Z_p is the metallicity of the stellar particle.

The contribution to the luminosity from the LMXBs from each stellar particle p is found using Eq. (4) in Fragos et al. (2013a), also with updated coefficients $\gamma_{(0-4)}$ from Madau & Fragos (2017),

$$\begin{aligned}\log\left(L_p^{\mathrm{LMXB}}/M_p\right) &= \gamma_0 + \gamma_1 \log(t_p/\mathrm{Gyr}) + \gamma_2 \log(t_p/\mathrm{Gyr})^2 \\ &+ \gamma_3 \log(t_p/\mathrm{Gyr})^3 + \gamma_4 \log(t_p/\mathrm{Gyr})^4 \\ &\times \left(\mathrm{erg\,s^{-1}\,10^{10}M_\odot}\right), \quad \text{for } t_p \in [0, 13.7]\,\mathrm{Gyr}\end{aligned} \tag{2.3}$$

where t_p (in Gyr) and M_p (in $10^{10}\mathrm{M}_\odot$) are the age and mass of the stellar particle, respectively.

The particles luminosities and spectral shapes are added as described in the previous section to obtain the corresponding source characteristics. Finally note that we adopt coefficients and spectral shapes that are not attenuated from interstellar absorption.

Thermal X-rays from hot ISM

We also include ionizing radiation from the diffuse ISM of galaxies. The spectral shape $\hat{S}_i^{\mathrm{ISM}}(\nu)$ is assumed to be that of thermal bremsstrahlung and constant in redshift (Pacucci et al., 2014),

$$\hat{S}_i^{\mathrm{ISM}}(\nu) = \begin{cases} C & \text{for } h_{\mathrm{P}}\nu \leq kT_{\mathrm{ISM}}, \\ C\left(h_{\mathrm{P}}\nu/kT_{\mathrm{ISM}}\right)^{-3} & \text{for } h_{\mathrm{P}}\nu > kT_{\mathrm{ISM}}, \end{cases} \tag{2.4}$$

where C is a normalization constant that ensures the correct units of Hz^{-1}, and kT_{ISM} is the thermal energy of the ISM in eV. From the spectral analysis of observations of the diffuse gas in galaxies by Mineo et al. (2012b), we use $kT_{\mathrm{ISM}} = 240$ eV, translating into a characteristic temperature of the heated ISM of $\sim 10^6$ K.

Each halo i has an individual luminosity $L_i^{\rm ISM}$, which we evaluate using Eq. (3) in Mineo et al. (2012b):

$$L_i^{\rm ISM}(0.3-10\,{\rm keV})/{\rm SFR}_i = (7.3\pm1.3)\times10^{39} \times\left({\rm erg\,s^{-1}M_\odot^{-1}yr}\right), \quad (2.5)$$

where 0.3-10 keV indicates the photon energy range this relation was obtained for. Note that $L_i^{\rm ISM}$ is corrected to be free of interstellar attenuation. We also rescale $L_i^{\rm ISM}$ to match our frequency band, which has the lower and upper limits of 13.6 eV and 2 keV, respectively.

Accreting nuclear black holes

To account for the ionizing photons originated from accretion disks surrounding nuclear black holes, we identify black hole particles in MBII and use their accretion rates to determine the production of ionizing photons. The bolometric luminosity of a black hole i is (Shakura & Sunyaev, 1973):

$$L_i^{\rm BH} = \eta\dot{M}_i c^2, \quad (2.6)$$

where η is an efficiency parameter, $\dot{M}_i$ is the accretion rate and c is the speed of light. Consistent with the black hole evolution and feedback in MBII, we choose $\eta = 0.1$.

As a spectral shape, we adopt the observationally derived mean QSO SED of Krawczyk et al. (2013), which is based on 108,104 QSOs sampled at $0.064 < z < 5.46$. When no observational data is available between 13.6 eV and 200 eV, this is derived as interpolation between the mean SEDs for which they have sufficient observations at both higher and lower energies. For energies greater than 200 eV, the spectral shape is modeled as a power law,

$$\hat{S}_i^{\rm BH}\left(h_{\rm P}\nu > 200\,{\rm eV}\right) \propto \nu^{-1}. \quad (2.7)$$

Finally note that no evolution of the SED with redshift is assumed.

Table 2.1: Thermal and ionization state of the IGM at $z = 14$, 12 and 10 for different combinations of source types. Note that ionization fractions below 10^{-5} are denoted with "<".

Source type	**T**[a] **[K]**				$x_{\mathbf{HII}}$			$x_{\mathbf{HeIII}}$		
	Median	Volume avg.	Mass avg.	Neutral avg.	Median	Volume avg.	Mass avg.	Median	Volume avg.	Mass avg.
$z = 14$										
Stars	10	16	23	16	<	7.144×10^{-5}	1.514×10^{-4}	<	<	<
Stars, BHs	10	16	23	16	<	7.146×10^{-5}	1.514×10^{-4}	<	<	<
Stars, XRBs	10	16	23	16	<	7.165×10^{-5}	1.516×10^{-4}	<	<	<
Stars, ISM	10	16	23	16	<	7.201×10^{-5}	1.522×10^{-4}	<	<	<
Stars, BHs, XRBs, ISM	10	16	23	16	<	7.217×10^{-5}	1.523×10^{-4}	<	<	<
Stars, BHs, XRBs, ISM (X=1, Y=0)	10	17	23	16	<	7.273×10^{-5}	1.533×10^{-4}	–	–	–
Stars, BHs, XRBs, ISM (X=0.92, Y=0)	10	17	24	16	<	8.304×10^{-5}	1.720×10^{-4}	–	–	–
$z = 12$										
Stars	10	73	142	51	<	1.627×10^{-3}	3.456×10^{-3}	<	<	<
Stars, BHs	10	73	142	51	<	1.630×10^{-3}	3.459×10^{-3}	<	<	<
Stars, XRBs	10	73	142	51	<	1.633×10^{-3}	3.462×10^{-3}	<	<	<
Stars, ISM	11	74	143	52	<	1.644×10^{-3}	3.478×10^{-3}	<	<	<
Stars, BHs, XRBs, ISM	11	74	143	52	1.004×10^{-5}	1.652×10^{-3}	3.486×10^{-3}	<	<	<
Stars, BHs, XRBs, ISM (X=1, Y=0)	11	75	142	54	1.381×10^{-5}	1.688×10^{-3}	3.545×10^{-3}	–	–	–
Stars, BHs, XRBs, ISM (X=0.92, Y=0)	11	79	149	56	1.524×10^{-5}	1.923×10^{-3}	3.960×10^{-3}	–	–	–
$z = 10$										
Stars	11	496	886	214	<	2.009×10^{-2}	3.581×10^{-2}	<	<	1.073×10^{-5}
Stars, BHs	11	497	887	214	<	2.012×10^{-2}	3.584×10^{-2}	<	<	1.238×10^{-5}
Stars, XRBs	12	498	888	215	4.442×10^{-5}	2.014×10^{-2}	3.586×10^{-2}	<	<	1.311×10^{-5}
Stars, ISM	16	504	897	219	1.008×10^{-4}	2.028×10^{-2}	3.604×10^{-2}	<	2.915×10^{-5}	8.005×10^{-5}
Stars, BHs, XRBs, ISM	18	506	900	221	1.464×10^{-4}	2.036×10^{-2}	3.612×10^{-2}	<	3.092×10^{-5}	8.408×10^{-5}
Stars, BHs, XRBs, ISM (X=1, Y=0)	19	510	876	244	1.927×10^{-4}	2.073×10^{-2}	3.664×10^{-2}	–	–	–
Stars, BHs, XRBs, ISM (X=0.92, Y=0)	19	558	941	252	2.111×10^{-4}	2.374×10^{-2}	4.100×10^{-2}	–	–	–

[a] The neutral average is calculated by weighting the values by the neutral fraction $x_{\rm HI}$ of the cell. More neutral cells then have larger impact. This quantity is relevant for studies of the 21 cm signal.

2.3 Results

A central question of this study is: can any of our source types drive an Epoch of Heating separate from an Epoch of Reionization? In other words, *is there heating preceding substantial ionization*? To examine this, we first discuss the qualitative and quantitative (in terms of average physical quantities) impact of the different source types on the IGM and its ionization and temperature histories (§2.3.1). Second, we focus on the thermal and ionization state of the IGM at $z = 12$ and 10, to discuss the details of these processes by using global maps, difference maps and phase diagrams (§2.3.2).

For the sake of clarity, throughout the manuscript we will use the following nomenclature to define the various ionization states of the IGM: *neutral* when $x_{\rm HII} < 10^{-5}$, *partially ionized* when $10^{-5} \leq x_{\rm HII} < 10^{-1}$, *highly ionized* when $10^{-1} \leq x_{\rm HII} < 0.9$, and *fully ionized* when $x_{\rm HII} \geq 0.9$.

2.3.1 Global history

Average ionization fractions and temperatures

In this section we present both mass and volume average values of various physical quantities, which are summarized in Table 2.1.

First, we turn to the **ionization fractions** of hydrogen and helium. We limit our discussion to $x_{\rm HII}$ and $x_{\rm HeIII}$, as the behaviour of $x_{\rm HeII}$ is very similar to the one of $x_{\rm HeII}$. From $z = 18$ to $z = 10$, they increase logarithmically with redshift, and the main driver is the stars, which dominate the emissivity budget. The scarcity of black holes make them irrelevant on cosmic scales at $z \geq 10$. Similarly, the other source types, albeit being widespread, do not contribute significantly to the statistics as they are far less luminous than stars. At $z = 18$, the volume as well as mass averaged ionization fractions are well below our convergence limits of 10^{-5}, independently from the combination of source types. This is true also for the median value (which provides insight into the state of most of the IGM, as also discussed by Ross et al. 2017), as the vast majority of the IGM is fully neutral. At $z = 14$, the first redshift shown in Table 2.1, the median value is still below 10^{-5} for all combinations of source types, while the volume and mass averaged ionization fractions now have increased to $\sim 10^{-5}$ and 10^{-4}, respectively. At $z = 12$ (i.e. 70 Myr later) the median $x_{\rm HII}$ is still below 10^{-5}, except for the case when all sources are present, for which it is 1.0×10^{-5}. The volume and mass averaged $x_{\rm HII}$ are higher ($\sim 10^{-3}$), with slight variations depending on the combination of source types. At $z = 12$, when all source types are present, we thus observe that the IGM is in the beginning stages of a cosmic, partial ionization phase.

Another 100 Myr later, at $z = 10$, the IGM has been under the influence of a factor of five more stellar sources, as well as one additional black hole. The ionization fractions have increased by an order of magnitude, reaching a volume average $x_{\rm HII} \sim 2 \times 10^{-2}$ independent of source types, and a mass average that is about twice as large. The median hydrogen ionization fraction ranges from below 10^{-5} (with stars or BHs), to 4×10^{-5} with XRBs,

to $> 10^{-4}$ with the ISM. We can understand the small differences in ionization fractions to come from their sensitivity to highly or fully ionized cells, whose prominence is totally dominated by the contribution from stars. The median, on the other hand is not as prone to biasing by the highly/fully ionized cells. When the ISM is present in addition to the stars, helium is also doubly ionized on cosmic scales, albeit only weakly. At $z = 10$ the IGM is predominantly fully neutral when only stars and black holes are present, while the majority of the gas is partially ionized in the presence of more numerous and diffuse energetic sources.

We now investigate the **temperature** evolution. At $z = 18$ the volume and mass averaged temperatures are 11 K and 12 K, respectively, irrespective of combination of source types. The median temperature is also comparable, being 10 K. The above statistics are identical to those we obtain directly from the hydrodynamical simulation MBII, i.e. they are unaffected by photo-ionization. At $z = 12$, the temperature statistics differ. The volume (mass) average is $\sim$ 70 (140) K, independent of the combination of source types, compared to a gas temperature of 38 (78) K in MBII. Weighing the temperature by the hydrogen neutral fraction we find a value $\sim$ 50 K. The small differences between combinations of source types indicate that stars are the main driver of the evolution of the averaged temperatures, as they dominate the global emissivity budget. The median temperatures however, are much lower, around 10 K. Having both XRBs and the ISM present in addition to stars and black holes does not raise the median temperature by $z = 12$.

At $z = 10$, we have larger differences in the temperature statistics. This coincides with the widespread partial (but low) ionization we found. The volume, mass and neutral fraction averaged temperatures are now in the range (496—506) K, (886—900) K and (214–221) K, respectively, depending on the increasing combination of source types. The median temperature is however the least biased indicator of the state of the majority of the IGM. With only stars (and BHs), it is 11 K, but with all source types present, it is 18 K. This is an order of magnitude higher than what we obtain from the MBII, where the median temperature at $z = 10$ is 6 K.

To summarize, we find that the IGM at $z = 12$ is showing its first signs of transitioning into one of two possible states, depending on source populations. With the most conservative assumption, i.e. in the presence of only stars (and black holes), the IGM is mostly fully neutral and cold (10 K), and it remains such down to $z = 10$. On the other hand, if XRBs and the radiation from the ISM in galaxies are also accounted for, the IGM becomes mainly partially ionized by $z = 10$. This very low partial ionization is accompanied by a further temperature increase of $\sim$ 10 K in all statistics. However, they also show that different regions of the IGM will have different temperatures. This in turn can have observable consequences. Thus, there is a modest Epoch of Heating, and it largely coincides with large-scale low partial ionization from energetic sources. We now turn to examine the details of this epoch.

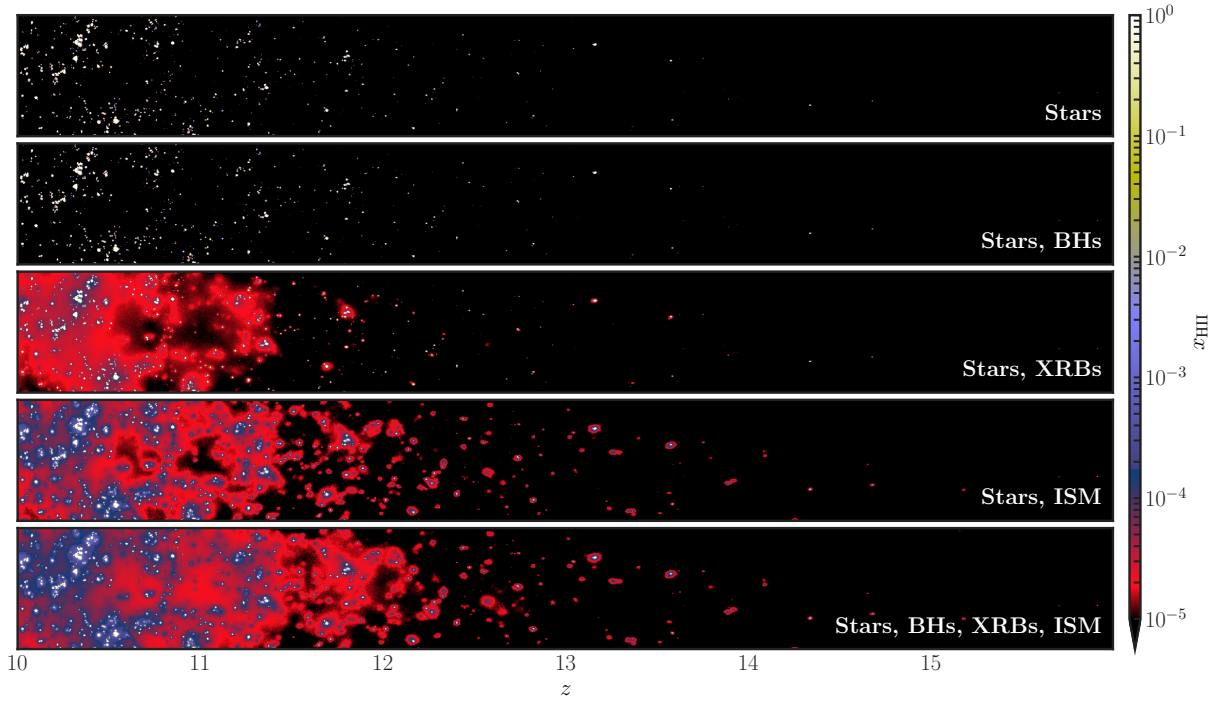

Figure 2.3: Lightcones showing the evolution of the ionized hydrogen fraction $x_{\rm HII}$ in the simulations with different combinations of source types, as indicated in the labels. The vertical size is $100h^{-1}$ cMpc and the aspect ratio between the axes is given at $z = 10$, hence the evolution in the angular diameter distance is not taken into account. The combined effect of all source types leaves the IGM without fully neutral regions at $z \lesssim 11.5$.

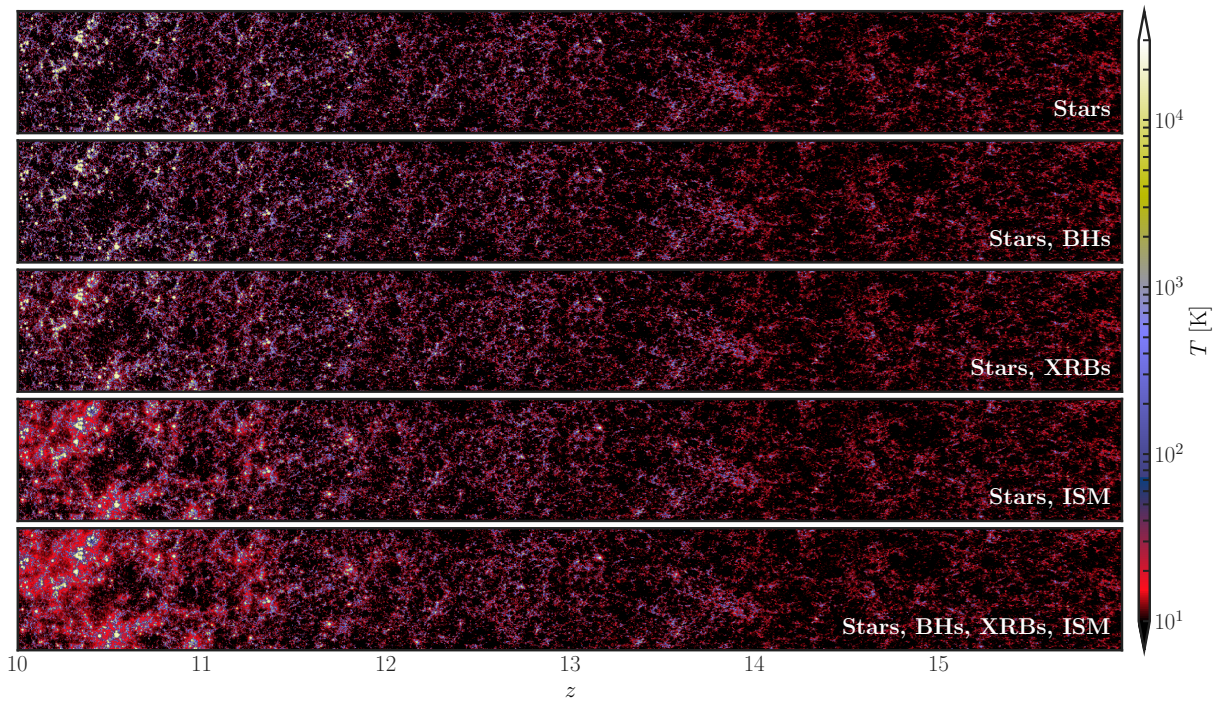

Figure 2.4: Lightcones showing the evolution of the IGM gas temperature under the presence of different source types as explained in Fig. 2.3.

Evolution on lightcones

To better understand the morphological impact of the source types, in Fig. 2.3 we illustrate how the hydrogen ionization state, $x_{\rm HII}$, evolves in lightcones from $z = 16$ to $z = 10$.

We find that stars are determining the existence and shape of the fully ionized regions: clusters of galaxies produce ionization bubbles with sharp ionization fronts. These bubbles grow mostly separately in our redshift range. At $z = 10$, the IGM is still predominantly neutral, but ionized bubbles are present within distances of tens of cMpc between each other. With stars we thus have a clear dichotomy with fully ionized bubbles residing in an otherwise fully neutral IGM.

As with the seeding procedure adopted in the MBII simulation only a handful of black holes is present at $z > 10$, the global history is unaffected by them. The first BH appears at $z = 13$, while the brightest is at $z = 11$ with $M_{\rm AB} = -17.6$. Its ionizing photon production rate is $\varepsilon^{\rm BH} = 2.05 \times 10^{53}$ phots s^{-1}, its mass is $M_{\rm BH} = 1.17 \times 10^6 h^{-1}{\rm M}_\odot$, and it accretes with a rate of $\dot{M}_{\rm BH} = 10^{-2}{\rm M}_\odot$ yr^{-1}. The stars within its host galaxy are even brighter, with $M_{\rm AB} = -18.7$, and produce $\varepsilon^{\rm stars} = 3.22 \times 10^{53}$ phots s^{-1}. In chapter 3 focused on the EoR we investigate the BHs' impact on the later stages of cosmic reionization, as well as the effect of a different seeding procedure.

Including more energetic galactic sources (XRBs and the ISM), we note a significant change in the ionization history of the IGM, which becomes partially ionized ($10^{-5} <$

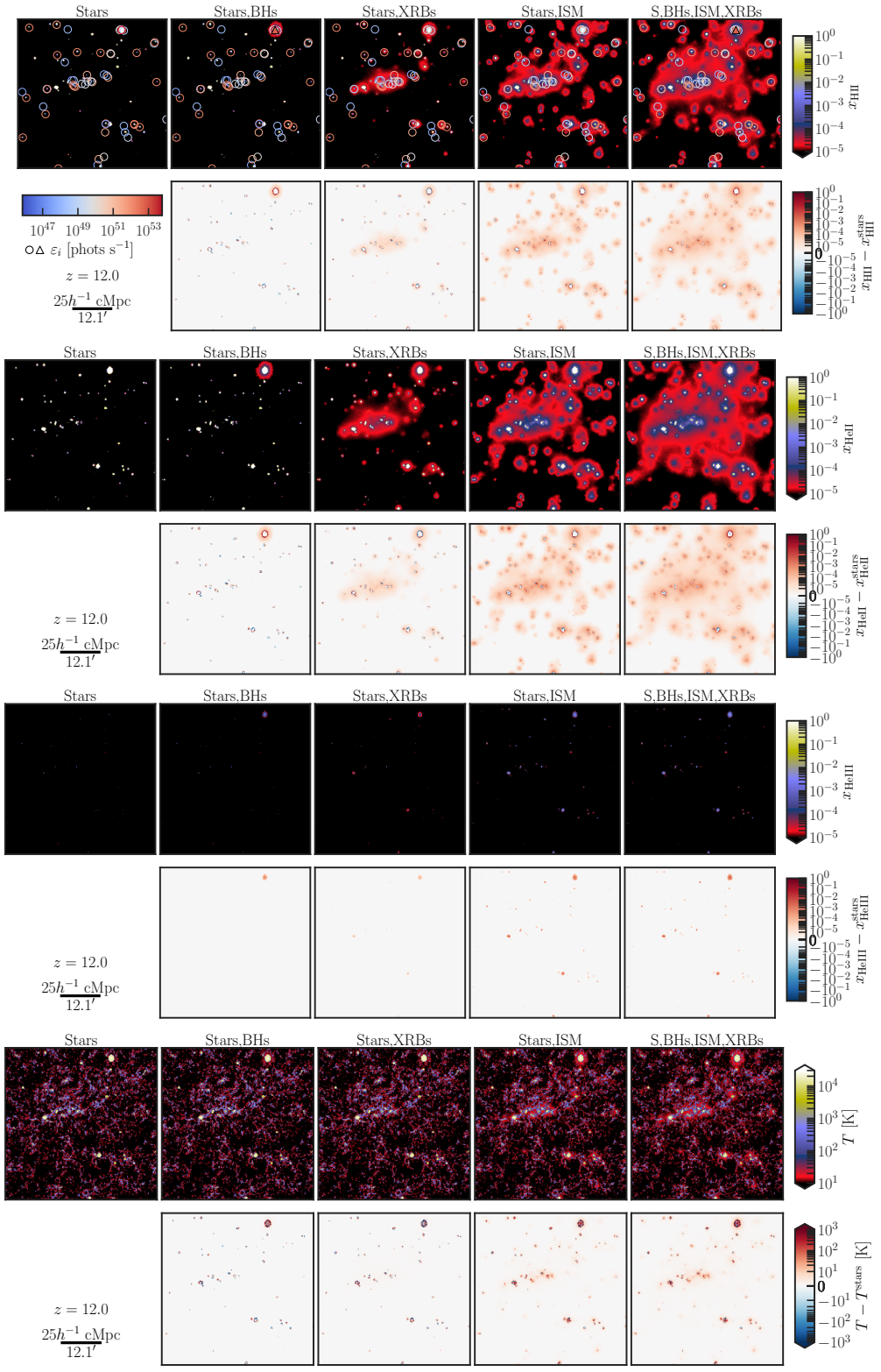

Figure 2.5: Maps (slices through the volume) in the plane of the only black hole at $z = 12$. The columns indicate different combinations of source types. The upper rows show the quantity in question, and the lower rows show the difference with respect to the first panel, which refer to simulations with stars only. In the maps for $x_{\rm HII}$, we also show the location of the sources as circles and their emissivities (denoted with the color). The only black hole is shown as a triangle in the upper right corner. The maps are $100h^{-1}$ cMpc wide.

$x_{\rm HII} < 10^{-1}$) several cMpc outside the fully ionized regions. On the other hand, these sources do not significantly contribute to extending the fully ionized regions, which are mainly governed by stars. The spectra of the XRBs peak at ~ 2 keV (Madau & Fragos, 2017), leaving fewer photons nearer the HI ionization threshold. Their long mean free path means they interact after being redshifted significantly, leaving smooth partial signatures of the order $x_{\rm HII} \sim 10^{-5}$ tens of cMpc outside their sources. The spectrum of the ISM is a broken power-law, with a high energy tail fainter than that of the XRBs, but more photons in the hard UV regime, which are more readily absorbed closer to the source. With the ISM, we thus see stronger partial ionization gradients, $x_{\rm HII} \sim 10^{-3}$–10^{-5}, and a more patchy ionization morphology compared to XRBs.

The combined effect of having all source types present in galaxies is shown in the lowermost panel of Fig. 2.3. The stars determine the extent of the *fully* ionized regions which, at $z = 10$ still make up a very small part of the IGM. The extent of the *partially* ionized regions goes well beyond those provided by XRBs and the ISM alone. Combined, they leave a much larger fraction of the IGM in a partially ionized state.

In Fig. 2.4, we show the temperature evolution lightcones as well. The stars effectively determine the temperature of 10^4 K in the regions we recognize from the previous plot to be fully ionized. With additional source types, the temperature of these regions does not markedly change, but we do however have heating of the otherwise cold, $T < 10$ K, IGM when it is subject to partial and low ionization. Combined, the XRBs and the ISM provide heating that extends further than either can provide alone, indicating that their combined effect is not simply dominated by the largest of each. At $z = 10$, there is thus a significant difference between the thermal states the IGM can be in, depending on the source types.

2.3.2 The IGM at $z = 12$ and $z = 10$

As we have found that the IGM may transition into a predominantly partially ionized, lightly heated state by $z = 10$ under the influence of XRBs and the ISM, here we examine its state at $z = 12$, when this transition has begun. This also coincides with the redshift at which we can investigate the effects of the first black hole in our simulation volume, and allows us to quantify its impact, if any. The reader can refer to Table 2.1 for some numbers.

Morphology

In Fig. 2.5 we plot maps of (from top to bottom) $x_{\rm HII}$, $x_{\rm HeII}$, $x_{\rm HeIII}$ and temperature at $z = 12$ for different combinations of source types. These are slices through the volume, chosen to contain the only black hole in the simulation box, residing in a bright galaxy in the upper right corner. This black hole accretes with a rate of $10^{-2}{\rm M}_\odot$ yr^{-1} and has a mass of $5.1h^{-1} \times 10^5{\rm M}_\odot$. Its absolute magnitude is $M_{\rm AB} = -16.2$ and it has an ionizing emissivity of $\varepsilon^{\rm BH} = 5.6 \times 10^{52}$ phots s^{-1}. It resides in a galaxy where the stars have a higher emissivity, $\varepsilon^{\rm S} = 4.6 \times 10^{54}$ phots s^{-1}, while the XRBs contribute with $\varepsilon^{\rm XRBs} = 1.4 \times 10^{50}$ phots s^{-1} and the ISM with $\varepsilon^{\rm ISM} = 4.1 \times 10^{51}$ phots s^{-1}.

We also indicate the position and emissivity of the sources with colored circles in the maps of $x_{\rm HII}$, while the black hole is represented by a filled triangle. To highlight the impact of different source types, we also plot absolute differences of the quantities with respect to the simulations with stars only. Note that we create the difference maps by truncating values below our numerical convergence limit of $x_i = 10^{-5}$ for i = HII, HeII, HeIII (see Appendix A.1).

The stars are responsible for the fully ionized HII bubbles and their morphology, as we also saw in the lightcones. HeII follows HII due to the similar ionization potentials, but as HeI has an ionization cross section progressively larger than that of HI at photon energies $h_{\rm P}\nu \geq 24.6$ eV, it is then more readily ionized, and its spatial distribution is generally more extended. Also note that sources having in their spectra higher energy photons (see the cases Stars+XRBs and Stars+ISM), which are redshifted at the simulated scales, provide a contribution first to HeI ionization and then to HI. These combined effects give us a larger extent of partially ionized $x_{\rm HeII}$ than $x_{\rm HII}$, while there is still little overlap between the fully ionized regions, leaving a very patchy HII and HeII ionization morphology. Where sources are more grouped together, they leave imprints on their surroundings similar to what one would expect of separate, but much brighter sources. The extent of ionization must thus be considered an effect of the ionizing photons arising within a volume, rather than from a single source. This helps the partially ionized regions to grow considerably in the cases with XRBs or the ISM.

The only BH at $z = 12$ accounts for partial ionization in its vicinity, spanning several cMpc, beginning with a smoothing of the otherwise sharp ionization front already created by its host and surrounding galaxies. However, a partially ionized tail is not an unique feature of BHs, as it is also seen outside galaxies that have XRBs and the ISM, which both contribute with high energy photons.

As in this epoch, galaxies are more common than BHs, their XRBs and ISM completely dominate the partial ionization at cosmic scales, while the contribution by the ISM component is certainly more ubiquitous than that from the XRBs. The last panels of both HII and HeII cases finally show the concerted impact of all sources, resulting in a larger overlap of red areas (HeII, in particular, is found in the process of merging into a single one) and in a smoother transition from the fully ionized to the neutral regions. HeII ionization, on the other hand, is much more sensitive to high-energy UV photons ($54.4 \leq h_p\nu \leq 200$ eV), while its cross section at soft X-ray energies decreases by 2 orders of magnitude. This is the reason why partially ionized HeIII regions with ionizations up to $x_{\rm HeIII} \sim 10^{-3}$ are found only within the ionized bubble of the BH or correlating with the brightest galactic sources: the strongest XRBs, and the strongest/most clustered galaxies with a hot ISM contribution.

Maps of the IGM temperature resulting from the various source combinations are finally found in the bottom panels. The thermal state of the IGM shows signatures both from the hydrodynamic simulation and the heating from the sources. The ideal gas of the hydrodynamic simulations is heated in overdense regions, whereas shocks occur at scales below the resolution of our grid ($\sim 400h^{-1}$ ckpc).

As fully ionized HII and HeII regions are mainly driven by stars, gas at photo-ionization

temperatures of $\sim 10^4$ K strictly correlates with them, as shown in the bottom left panel by white areas. Extended blue areas, corresponding to temperatures of $\sim$ 500 K, are found along filaments connecting the stellar sources. These regions, on the other hand, do not have a clear counterpart in the hydrogen ionization pattern because their ionization fraction is below the convergence threshold $x_i < 10^{-5}$.

Difference maps show that the addition of other source types does not significantly change the global temperature pattern as setup by the stars. A local increase in T (from 100 up to 1000 K) is found only around the BH and around galaxies having a hot ISM. Interestingly, we note that source clustering still plays a relevant role in changing T but its pattern is less extended compared to the one of the corresponding HII region. Also note that the temperature boost expected in the HeIII regions from the release of HeII electrons, as found in Kakiichi et al. (2017), cannot be appreciated around the first quasars because of the lower emissivity and softer spectral shape. In fact, their BH had an emissivity $\varepsilon^{\rm BH} = 1.4 \times 10^{56}$ phots s^{-1}, and a power law spectrum with spectral index -1.5, shifting more ionizing photons closer to the ionization potential of HeII.

By comparing the relative contribution of different sources we immediately see that XRBs do not heat appreciably the IGM which they partially ionize. We find instead that the ISM emission, mainly providing softer photons, contributes to a heating of some tens of degrees in the regions that it partially ionizes. These do not extend more than a few cMpc outside the fully ionized regions. The different impact of XRBs and ISM can be easily explained in terms of their spectral distributions (see Fig. 2.2). The energy released by an absorbed X-ray photon is in fact shared between direct photo-ionization, secondary ionization and gas heating. The fraction of the energy that goes into heating increases with the ionization fraction of the gas. Heating is more effective for less energetic photons (Dalgarno et al., 1999), and the ISM, which has a softer spectrum, is therefore more effective in heating the gas.

Combining the effect of XRBs and the ISM, we expect a more efficient heating: XRBs should in fact drive pre-ionization in regions that can be influenced by successive contribution of ISM photons. This concerted effect is clearly visible as an increase of the extent of the heated regions created by the presence of all source types. Also check the difference map to have a better feeling of the difference introduced by the concerted, self-consistent, impact of all the source types.

To summarize this section: our models predict an IGM at $z = 12$ that is dichotomous, with many, patchy fully ionized and hot regions, which extent and abundance are determined by the stars. These regions reside in an IGM that is either neutral and cold, or partially ionized when XRBs and the hot ISM of the galaxies irradiate harder UV and X-ray photons. They can heat the IGM only up to a few tens of degrees, and although this heating is certainly boosted when the XRBs pre-ionize the IGM, it is the ISM that sustains the heating.

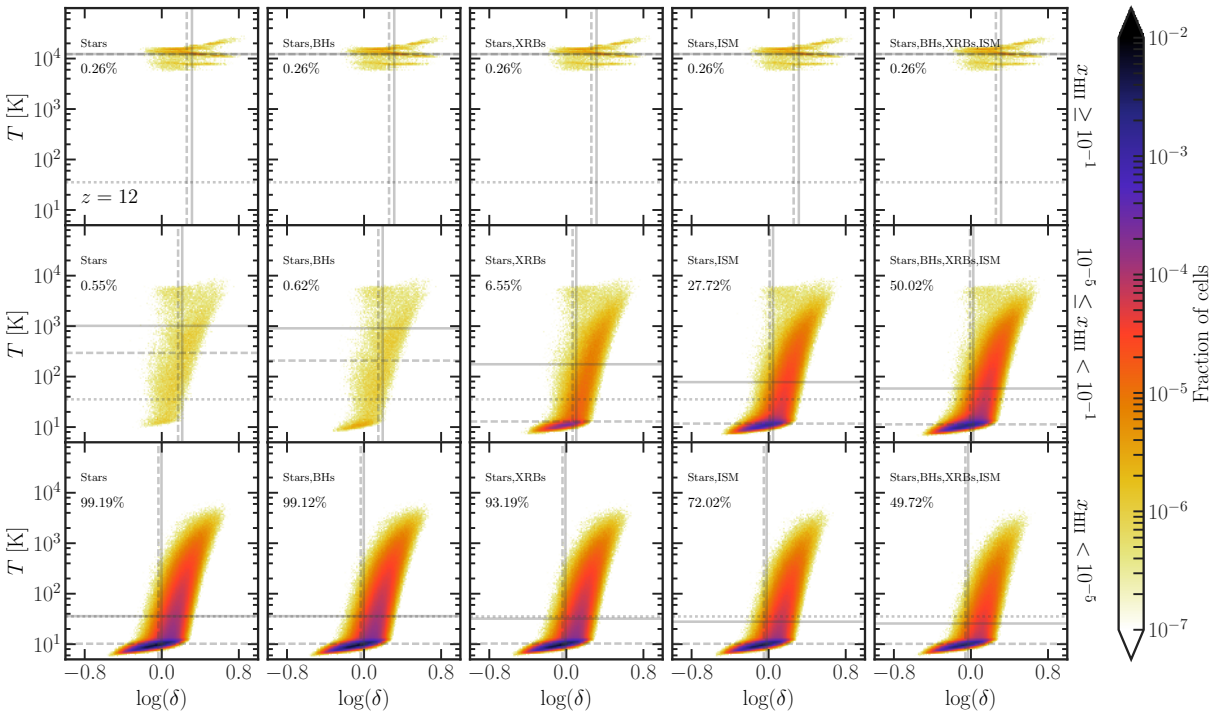

Figure 2.6: Phase diagrams showing the distribution (in fraction of the total $256^3 = 1.68 \times 10^7$ cells) of the IGM at $z = 12$ in different thermal (T, y-axes), overdensity ($\delta \equiv n/\bar{n}$, x-axis) and ionization ($x_{\rm HII}$, rows) states for various combinations of source types, as indicated by the labels. The horizontal solid (dashed) lines indicate the mean (median) temperature in the volume, and the vertical solid (dashed) lines indicate the mean (median) overdensity δ of the part of the volume with the given ionization state. The horizontal dotted lines indicate the CMB temperature. The percentages indicate the fraction of the volume that is in the given ionization state.

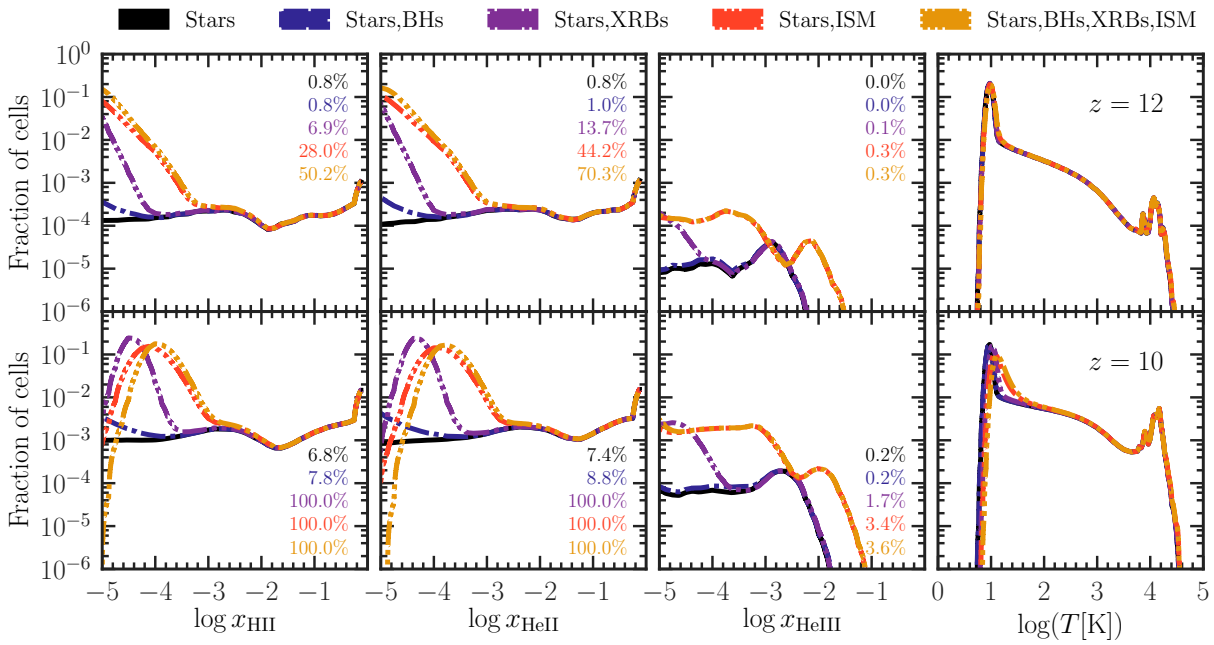

Figure 2.7: Fraction of cells in different ionization and thermal states for different combinations of source types (*stars*: solid black, no ticks; *stars* and *BHs*: blue, one tick; *stars* and *XRBs*: purple, two ticks; *stars* and *ISM*: red, three ticks; *all*: yellow, four ticks). We show the states at $z = 12$ (upper row) and at $z = 10$ (lower row). The percentages indicate the fraction of the IGM which is shown in the histogram. The remaining IGM has an ionization fraction x_i lower than 10^{-5} (for $i =$ HII, HeII, HeIII).

IGM phase statistics

In Fig. 2.6 we show the phase state of the IGM at $z = 12$ under the influence of different combinations of source types. Each panel refers to the distributions of the total 1.68×10^7 cells with a given temperature as a function of overdensity δ[4].

The upper row shows the highly and fully ionized IGM to be independent from the type of sources. The IGM in this state is hot, with mean and median temperatures of 1.2×10^4 K at a median overdensity $\delta = 2$. The stars that drive full ionization primarily affect the overdense regions. As reflected in the values shown in Table 2.1, a negligible fraction of the gas is in this high ionization state at $z = 12$.

The middle row shows the phase state of cells with partially ionized IGM. We see that a very small fraction is in this state with stars alone, and adding a single BH does not change this. Including XRBs (ISM) increases the percentage of volume in this partially ionized state to 7% (28%), while having all sources present, 50% of the volume becomes partially ionized. A comparison with the lower row, which refers to the neutral IGM, clearly show the effect of the additional energetic photons. The IGM, under the presence of only stars is mostly cold, but with a fraction of cells that is both overdense ($\log \delta > 0$) and hotter than the majority ($T > 10$ K). A larger fraction of these cells transitions to a partially ionized state when including the XRBs. An even larger fraction, both from the cold and underdense as well as the overdense, hotter gas, becomes partially ionized under the influence of the ISM photons. There is no apparent preference in terms of δ or T on the IGM that becomes partially ionized. This results from a combination of factors such as the location (overdense regions), emissivity (this is lower for the XRBs than for the ISM) and spectral shape (the XRBs have spectra harder than the ISM) of the sources. More than 99% of the IGM is fully neutral with stars only, while this fraction is strongly decreased with XRBs (93%), the ISM (72%) or all source types (50%).

It should be noted that, as already pointed out by other authors (see e.g. Ross et al., 2017), the partially ionized and warm cells found in the presence of stellar type sources might only arise from a lack of spatial resolution. More specifically, whenever the cell size is too large to resolve the sharp ionization front expected from stellar type sources, the cell containing the front appears partially ionized and warm, while in reality part of the gas in the cell should be neutral and cold, and part fully ionized and hot.

In Fig. 2.7, we present histograms of the temperature and ionization fractions at $z = 12$ and $z = 10$, showing a quantitative evolution of the physical state of the IGM.

As a baseline, we compare the distributions to the case with stars alone, where the curves for $x_{\rm HII}$ and $x_{\rm HeII}$ are almost flat, with only a surplus of cells in a highly or fully ionized and hot ($T \sim 10^4$ K) state, while the large majority is neutral and cold ($T \sim 10$ K), both at $z = 12$ and $z = 10$. The fraction that is highly/fully ionized and hot has increased an order of magnitude by $z = 10$. With the addition of other source types, partial ionization is strongly increased, especially at $x_{\rm HII} < 10^{-3}$. Only at $z = 10$ a discernible difference is visible in the temperature distributions, with a shift of the low temperature peak by several

[4]The gas overdensity is in this chapter defined as $\delta \equiv n/\bar{n}$, where n is the particle number density with units cm^{-3}.

tens of degrees when all source types are present. This is also the redshift where there is largest difference in the fraction of the IGM that has $x_{\rm HII} > 10^{-5}$, denoted as percentages in the figures. At $z = 12$, 1% of the IGM has $x_{\rm HII} > 10^{-5}$ when only stars are present, while this fraction has increased to 50% with all sources. At $z = 10$, this difference is even more significant, as the percentage goes from 6% to 100% when including either XRBs, the ISM, or both. The behaviour of $x_{\rm HeII}$ closely follows the one of $x_{\rm HII}$.

The distribution and presence of doubly ionized helium changes significantly depending on the source types. While the presence of XRBs has an impact only at $x_{\rm HeIII} < 10^{-4}$, the ISM ionizes more and to higher values of $x_{\rm HeIII}$. At $z = 12$, less than 1% of the IGM has $x_{\rm HeIII} > 10^{-5}$. This fraction, as well as the distribution across all $x_{\rm HeIII}$, increases by an order of magnitude for all source types at $z = 10$, leaving 4% of the helium in the IGM in a partially ionized state.

2.3.3 Simulations without helium

The presence of helium in our simulations affects the ionization state of hydrogen as well as the temperature evolution. The ionization cross section of hydrogen decays roughly as ν^{-3}, and at the ionization threshold of 24.6 eV for HeI, the cross section of helium is a factor of 6 larger than that of hydrogen. Similarly, at 54.4 eV, the ionization potential of HeII, the helium cross section is 16 (neutral) and 13 (singly ionized) times larger than that of hydrogen. Sources that have spectra that allow for helium to get ionized twice will result in the release of free electrons that can further heat the IGM, as found by Ciardi et al. (2012).

We have run simulations without helium, i.e. $Y = 0$, including all source types. We have either replaced the now absent helium with hydrogen, effectively increasing its number fraction from $X = 0.92$ to $X = 1$, or we have simply not solved for helium chemistry, but kept using $X = 0.92$. Ionization fractions and temperature statistics are presented in Table 2.1. When $X = 0.92$, we effectively have more photons to ionize hydrogen atoms and heat the IGM, resulting in ionization fractions and temperatures higher at any given redshift compared to the case in which helium is included. This effect can be mitigated to some extent by increasing the number fraction of hydrogen to $X = 1$.

In Fig. 2.8, we have histograms showing the distribution of $x_{\rm HII}$ and T at $z = 12$ and 10. Dashed and dashed-dot lines indicate simulations without helium. The increase in ionization fractions mentioned above is accompanied by a uniform shift at $x_{\rm HII} < 10^{-1}$ towards higher $x_{\rm HII}$ at both z, while the distributions of highly and fully ionized cells do not change. The absence of helium also contributes to another interesting effect: the maximum temperature reached in the IGM is lowered by approximately 0.2 dex, Without helium, at $z = 12$ ($z = 10$) the maximum temperature is $T = 22,986$ $(36,141)$ K, while with helium it is $T = 32,313$ $(38,709)$ K. We also have some additional heating of the cold IGM at $z = 10$, as seen from the increase in the median temperatures in Table 2.1.

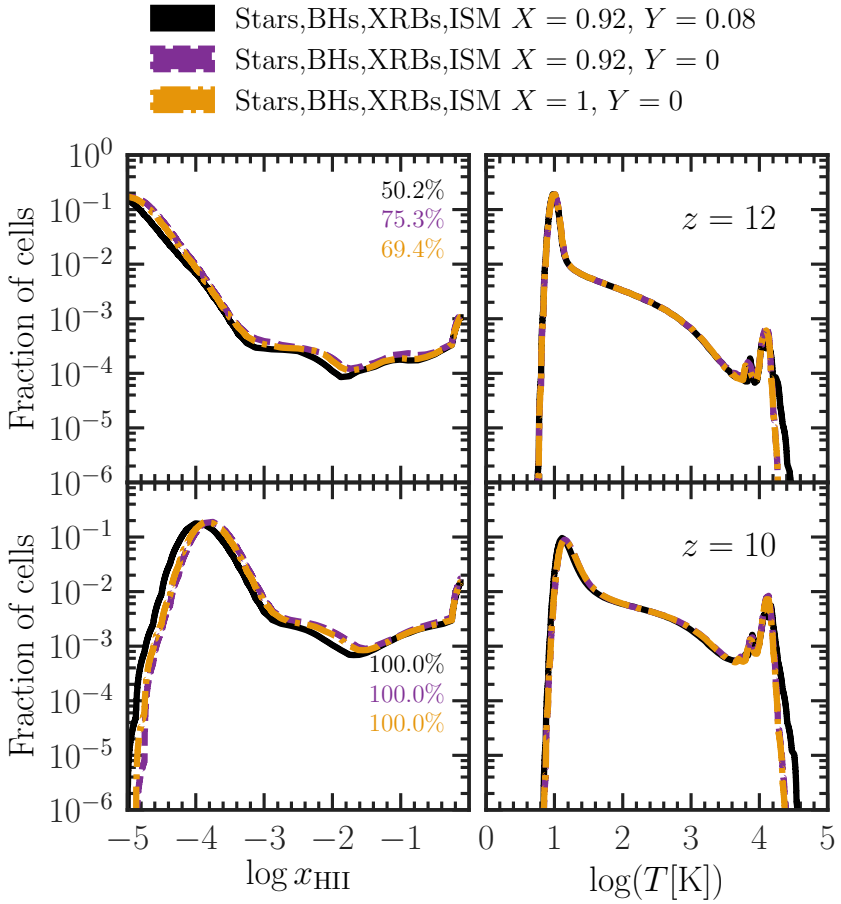

Figure 2.8: Distributions of volume averaged temperature and ionization fraction of hydrogen for simulations with (solid, black line) or without helium (dashed; purple lines are for $X = 0.92$ and $Y = 0$, while dash-dotted yellow lines are for $X = 1$ and $Y = 0$) at $z = 12$ and 10. The percentages, with the same colours and reference as the lines, indicate what fraction of the IGM has $x_{\rm HII} \geq 10^{-5}$, and hence is partially or strongly ionized.

2.4 Discussion

Our simulations show that stars are the principal source of IGM ionization throughout the EoH: in fact they fully dominate the UV part of our spectra (see Fig. 2.2) and have a large spatial coverage. They are responsible for the volume averaged ionization fractions of both hydrogen and helium, which at $z = 10$ are $x_{\rm HII} \sim x_{\rm HeII} \sim 2.0 \times 10^{-2}$. The contribution of stars also drives the shape and extent of local HII and HeII bubbles, having sharp ionization fronts due to the short mean free path of their UV photons. As a rapid phase transformation is not expected until $x_{\rm HII} \sim 0.1$ (Furlanetto & Oh, 2016), at $z = 10$ the IGM is still patchy, with its vast majority being fully neutral, and only a small fraction fully ionized. Only in the presence of more energetic sources, does the morphology and state of the ionized regions change.

The seeding procedure adopted in MassiveBlack-II, plants seeds of mass $10^5 {\rm M}_\odot$ in halos with $M > 10^{10} {\rm M}_\odot$. This results in a scarcity of BHs at $z > 10$, which makes their contribution to IGM ionization and heating negligible. This conclusion, though, is strongly model dependent. We discuss in more detail the effect of the seeding procedure and the possibility of populating with BHs also smaller halos in chapter 4.

It is also important to note that the signature of BHs is degenerate with that of other, more abundant, energetic sources. The X-ray binaries, accounting for both the sub-dominant low mass binaries (LMXBs) and the SFR-tracing high mass binaries (HMXBs), have spectra that peak at keV scales. They provide a smooth, long-range (several cMpc) partial ionization of hydrogen and helium ($x_{\rm HII} = x_{\rm HeII} \gtrsim 10^{-5}$). Otherwise, they contribute negligibly, in line with the findings of the semi-analytical works of Madau & Fragos (2017) and Sazonov & Khabibullin (2017). Our model of XRBs thus ionizes less than what found by Ross et al. (2017), who employ a comparable 3D RT post-processing approach, albeit on a dark matter N-body simulation. It should be noted though that their emissivities and source locations were obtained through halo mass scaling relations, which differ from our approach relying on a consistent evolution of baryonic sources (and their properties) from a hydrodynamical simulation. In particular, by $z = 13$ Ross et al. (2017) find a partial ionization of the cold IGM that is orders of magnitudes higher than what indicated by our computation. They also find a higher volume averaged $x_{\rm HII} \sim 10^{-2}$ independent of whether XRBs were included or not. However, while our XRB spectra peak at keV-scales, theirs is a power law with index -1.5. The luminosity of our XRBs scales with either the SFR and metallicity (HMXBs) or the stellar age and mass (LMXBs) in the halo where they reside, while Ross et al. (2017) assume a constant X-ray production efficiency of each halo. Our results are more in line with those of Meiksin et al. (2017), who found the HMXBs to heat the IGM to $T = 22$ K by $z = 10$. In their work, though, the increment in temperature from a model with stars only is ~ 20 K, compared to our ~ 1 K. This could be due to our RT approach capturing only the effect of X-rays up to distances of the order of the box length. On the other hand, Meiksin et al. (2017) do not consider photons with energies below 200 eV, which are more relevant on smaller scales. Note also that we are still likely to overestimate the already small contribution from XRBs. Das et al. (2017), for example, found the XRB spectra to be attenuated by the ISM, albeit not as strongly as advocated

by Fragos et al. (2013b), but still more than what we have assumed with our constant escape fraction of $f_{\rm esc} = 15\%$ for photons with energies lower than 200 eV.

Our multifrequency RT highlighted the importance of the ISM contribution as its spectrum provides large abundance of photons in the hard UV and soft X-rays. We found that this spectral difference is of great importance, a conclusion also shared by Pacucci et al. (2014), or by Fialkov et al. (2014b), who similarly showed that the hardness of the spectral shape of the XRBs would make them inefficient in heating the IGM. There is however some uncertainty on the extent of the contribution from the ISM, which could be larger, as the supernova-driven heating process of the ISM gas will likely result in emission outside the galactic planes (see e.g. Chevalier & Clegg 1985 or Strickland & Stevens 2000). In this case, it is less likely that the ionizing UV photons are strongly absorbed by the host galaxy, which is what we assume with our escape fraction. However, although the production rate of UV photons is uncertain, the ISM is known to supply galaxies ubiquitously with observable amounts of soft X-rays (Mineo et al., 2012b). The Meiksin et al. (2017) work mentioned above includes also the contribution of the ISM, which raises the IGM temperature at $z = 10$ to $T = 6$ K or $T = 34$ K (compared to our median $T = 16$ K with stars and the ISM), depending on the wind model adopted.

One should be careful to attribute long range partial ionization and heating signatures, if observed, solely to the ISM. These could possibly be degenerate with those from cosmic rays. While Leite et al. (2017) found that they should heat the IGM in the immediate vicinity of star forming halos, this happens because of the cosmic rays confinement. If this confinement is weaker than what estimated by Leite et al. (2017), their contribution could be very similar to that of the ISM.

Another RT study examining the effect of X-rays on heating and ionization is that of Baek et al. (2010). They include X-rays from all of their sources, labeling this contribution as QSOs, but noting that XRBs and supernova remnants also fall into this category. Their scaling of the X-ray luminosity with the SFR allows for some comparison to our case where all source types are present. Their mean ionization fraction at $z = 10$ is $x_{\rm HII} \sim 10^{-2}$ for simulations with helium, but without X-rays. They note that including QSOs that account for up to 10% of the total flux does not change this result, similarly to the findings of Ross et al. (2017), and now also ours. We have run simulations without helium, finding that the ionization fraction, contrary to Baek et al. (2010) (cf. their models S1 and S3) increases to $x_{\rm HII} = 2.4 \times 10^{-2}$ (keeping $X = 0.92$), or $x_{\rm HII} = 2.1 \times 10^{-2}$ (with $X = 1$). This is caused by the additional photons that are available for HI ionization when helium is absent, aligning our results with those of Ciardi et al. (2012). Including helium thus provides a slower progression of the cosmic heating and reionization. As for HeIII, we find that only faint, diffuse regions having $x_{\rm HeII} \sim 10^{-3}$ start to appear near bright energetic sources. We do not find the widespread existence of highly/fully ionized HeIII that Ross et al. (2017) found at higher redshifts.

Consistently with Madau & Fragos (2017), we find that XRBs are not able to heat significantly the IGM due to their hard spectra. Our results show that also the ISM, which has a much softer spectrum, is not able to raise the IGM temperature significantly by $z = 10$. Our median temperatures are in fact $T = 11$ K and $T = 18$ K for simulations

with stars only and all sources, respectively. This is somewhat higher than the value reported in Baek et al. (2010), where the temperature of the IGM with an ionization fraction $x_{\rm HII} < 10^{-2}$ was ranging between 3 K and 7 K without and with QSOs, respectively. Our volume averaged temperatures are, on the other hand, much lower than those of Ross et al. (2017), who found $T \sim 600$ K at $z = 13$ with stars only and $T \sim 1000$ K with HMXBs, compared to our volume averaged $T = 34$ K at $z = 13$ with stars and XRBs.

Predicting the thermal state of the IGM during the EoH is of crucial importance for 21 cm experiments because an IGM with a temperature higher than that of the CMB will be seen in emission, and if lower, it will be seen in absorption. Interestingly, at $z = 10$ the CMB temperature is $T_{\rm CMB} = 30$ K. The margins between a 21 cm signal seen in emission or absorption are thus close. Note, though, that while our mean temperatures at this z are much higher than $T_{\rm CMB}$, they do not represent the state of the vast majority of the IGM, which is either fully neutral (with stars only) or partially ionized (with all sources). We find partially ionized regions that are warmer than the CMB already at $z = 12$, as indicated from the neutral and mass averaged temperatures. We defer to a future work further investigation of the effect that different source types leave on the 21 cm signal.

Observationally, very late (after $z = 8.4$) heating of the IGM has been ruled out by measurements of 21 cm line (Pober et al., 2015). Less constrained is a period of early and highly efficient heating, which would suppress small scale structure formation by increasing the Jeans mass (Ostriker & Gnedin, 1996), disfavouring H_2 creation associated with X-ray fluxes that do not dominate the ionizing budget (Oh, 2001), and increase the Thomson-scattering optical depth of the CMB (Ricotti & Ostriker, 2004).

We observe that both the distribution of the sources and their surrounding ionized regions trace the underlying density field. This can be seen particularly well in the phase maps of Fig. 2.6, where the fully ionized regions reside where $\log \delta > 0$. We also see clustering of the sources in the global maps (Fig. 2.5). This clustering induces (a) ionized regions that are more extended and (b) heating of the nearby neutral IGM, when comparing the surroundings of sources with similar emissivities but that exist either in a clustered environment or alone. This distinction (i.e. clustered sources vs single source) is crucial in the interpretation of observations aiming at investigating the physical properties of the high redshift IGM.

2.5 Conclusions

In this chapter we have examined the sources that could end the *Dark Ages*, initiate the *Cosmic Dawn*, and drive the *Epoch of Heating* (EoH), when the intergalactic medium (IGM) starts transitioning from a cold and neutral phase, to a hot and ionized one. The Cosmic Dawn concludes with the *Epoch of Reionization* at $z < 10$, which we explore in chapter 3. We have investigated how *stars*, *X-ray binaries* (XRBs) and the *shock heated interstellar medium* (ISM) of galaxies, as well as *accreting nuclear black holes* (BHs), could heat and ionize the IGM at cosmic scales. Each source type has spectral characteristics determined by underlying physical processes. The luminosity of the stars is determined

by their masses, ages and metallicities, while the ISM and XRBs are sensitive to the star formation rate in galaxies, as well as their masses and metallicities. The black holes, on the other hand, shine with the rate they accrete matter.

We have used the hydrodynamical simulations MassiveBlack-II (MBII, Khandai et al., 2015) to provide us with the physical properties of the sources, their location and abundance, as well as their environment (temperature and baryonic density). Having established the sources, their spectra and ionizing emissivities, we post-processed MBII with the radiative transfer code `CRASH` (e.g. Ciardi et al., 2001; Graziani et al., 2013).

Our main findings can be summarized as follows.

(i) Stars drive the extent, shape, abundance and temperature of the fully ionized HII and HeII regions, but do not leave the remaining IGM in any heated or partially ionized state. Only 7% of the overall IGM is in a non neutral state at $z = 10$ with stars alone.

(ii) With our seeding prescription, nuclear black holes are scarce and do not contribute significantly to heating or ionization at $z > 10$.

(iii) XRBs contribute to partial, uniform ionization ($x_{\rm HII} \sim 10^{-5}$) of the dilute IGM on cMpc scales, but do not significantly heat it by $z = 10$. This can be attributed to their hard, keV-peaked spectra.

(iv) The ISM of the galaxies contribute to a larger extent of partial ionization than the XRBs do. Their softer spectra provide heating as well as ionization, with $x_{\rm HII} \sim 10^{-3}$ and $T \gtrsim 10$ K up to a few cMpc around the fully ionized regions at $z = 12$ and lower partial ionization further out.

(v) In concert, the ISM and the XRBs induce an an ionization fraction $x_{\rm HII} \geq 10^{-5}$ in 50% (100%) of the IGM at $z = 12$ ($z = 10$). However, at $z = 10$ the IGM is still predominantly cold, with median temperatures in the range (11–19) K under the influence of stars to all source types.

(vi) At $z = 10$ the IGM will be seen in both emission and absorption in 21 cm as the neutral averaged temperatures are $T > 200$ K for all combinations of source types.

(vii) Helium reionization has begun on cosmic scales at $z = 10$ in the presence of XRBs and/or the ISM.

If an Epoch of Heating takes place at $z > 10$, our findings indicate that it will be a modest one.

Chapter 3

The Epoch of Reionization

3.1 Introduction

As already discussed in chapter 2 in the context of Cosmic Dawn, there is a subtle interplay between the different sources of ionizing radiation, which is particularly relevant for a correct determination of the IGM temperature, but also for partially ionized H and He ionization. This, in turn, is expected to have an important impact, among others, on the modeling of the 21 cm signal from neutral hydrogen (see e.g. Ciardi & Madau 2003). Here, we thus push the simulations presented in chapter 2 and Eide et al. (2018) to a lower redshift and concentrate on the analysis of the effect of different sources of ionizing radiation at a later time, for $z \lesssim 10$. The results of this chapter are published in Eide et al. (2020a).

This chapter is structured as follows. In section 3.2 we present the details of MBII, the RT and the source modeling. In section 3.3 we present our results in terms of discussion of the physical state of the IGM as determined by different source types, as well as of comparison to available observations. We discuss the results and give our conclusions in section 3.4.

3.2 Method

The approach followed to model cosmic reionization is the same one presented in chapter 1 and chapter 2. We refer the reader to the mentioned chapters for more details.

In this chapter, we grid 15 snapshots from MBII[1] between $z = 18$ and 5 onto 256^3 cells, resulting in a resolution of $391h^{-1}$ ckpc. This gives us locations of potential ionizing sources through stellar particles, halos and black holes representative, respectively, of stellar populations, galaxies and galactic nuclei in various states of activity. We also grid the gas temperature and density[2].

[1]The redshift of the snapshots is dictated by the outputs of the MBII simulations and it is z=18, 16, 14, 13, 12, 11, 10, 9, 8, 7.5, 7, 6.5, 6, 5.5, 5.

[2]The gas density is translated to hydrogen and helium number densities assuming number fractions of

At $z < 10$ we further reduce the number of sources by evenly absorbing the luminosity of faint cells into all their neighbours that are at least 100 times brighter, if they themselves are not at least 100 times brighter than any neighbour. This procedure is iterated until no more cells can be absorbed into brighter neighbours. The result is a source field with a greatly reduced number of sources. The structure of this new field still closely resembles that of the original one, as only very faint extensions of larger source clusters are integrated into the brighter central cells that would outshine them in any case. Isolated faint sources are left untouched. The combination reduces the number of source by 50% at $z = 10$ and up to 70% at $z = 7$. Note that, with this procedure, at $z = 7$ the vast majority of sources are displaced by less than 1 h^{-1}cMpc and virtually all remain within 2 h^{-1}cMpc, i.e. much less than the typical dimension of an ionized region at the same redshift (see also Busch et al., 2020). As a reference, for the simulations with all sources included and $f_{\rm esc} = 15\%$ at $z = 8.56$ we find that clustering the sources induces a difference in the volume averaged temperature and HII and HeII fractions of 0.09%, which increases to 0.3% for the HeIII fraction. These values are within the numerical noise as they are below our convergence limits. A further comparison of cell by cell values indicates that the clustering of the sources does not affect the results presented in this chapter.

As in chapter 2, the emissivities of stars, XRBs and ISM that share the same cell are summed together and assigned a spectrum which is the sum of the volume averaged spectra of each source type. BHs are instead treated separately, as not every galaxy hosts an active BH. In Fig. 3.1 we show the spectral energy distribution (SED) of the various sources at different redshifts. Note that this is very similar to Fig. 2 of chapter 2 but has been nevertheless included to show the spectra at all redshift for completeness. As discussed in chapter 2, the evolution in z is extremely mild. The SEDs of the various sources are discussed in the following.

In Fig. 3.2 we show the luminosity function (LF) of the various source types at $z = 6$, estimated at an absolute magnitude $M_{\rm AB}$ at 1450 Å[3]. We note that the LF is calculated before applying any grid shifting of the sources and estimated on a 1024^3 grid, i.e. with sources grouped together in cells with widths of 143 ckpc, slightly larger than present-day Milky Way-sizes. The LFs are discussed in the following, where we summarize the main features of the four different adopted source types.

Stars: The spectra of stars are obtained from the 2012 version of the stellar population synthesis code `BPASS` (Eldridge & Stanway, 2012) by evaluating a spectrum for each star particle given its stellar mass, age and metallicity. We do not account for contributions from stars that have evolved into binary systems, i.e. we use the 'single star' version of BPASS, nor of nebular emission. As a reference, at $z = 6$ we have a total of 3.1×10^7 star particles. The volume averages of the individual galaxy spectra are strongest at HI and HeI ionizing frequencies, and fall by orders of magnitude near the HeII ionization threshold, as seen in Fig. 3.1. The spectra display little redshift evolution towards lower z except for a

$X = 0.92$ and $Y = 0.08$, respectively, and no metals.

[3]Except for XRBs, for which we estimate the LF at a wavelength of 6.2 Å, corresponding to photon energies of 2 keV.

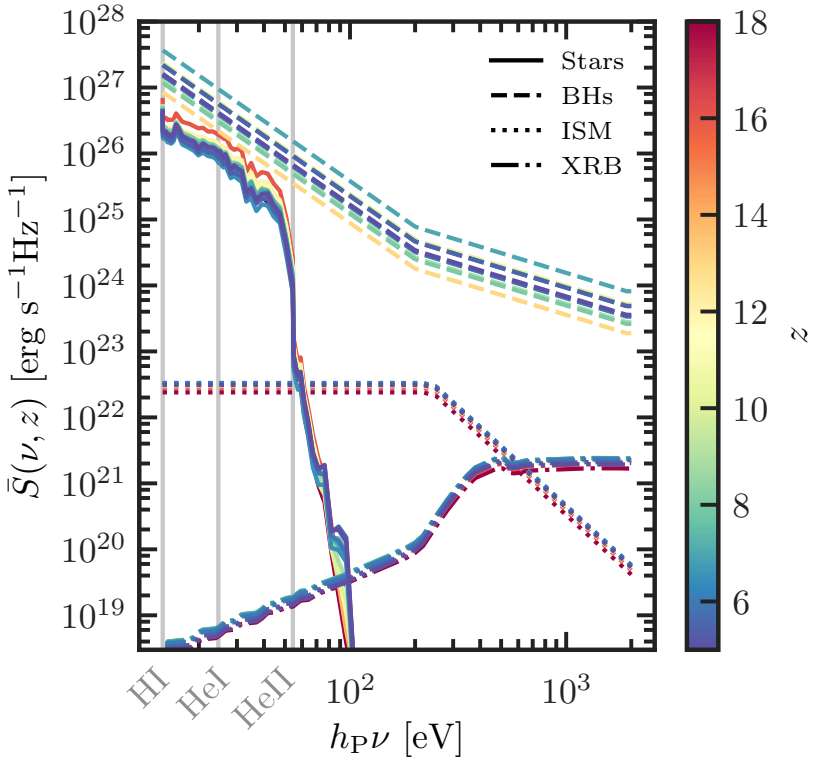

Figure 3.1: Averaged SEDs for each source type, stars (solid lines), BHs (dashed), ISM (dotted) and XRBs (dash dot-dotted) as function of photon energy plotted at various redshifts (indicated by the line colour). The ionization thresholds for HI, HeI and HeII are plotted as vertical grey lines.

dampening of the spectra at lower energies and a hardening at energies > 54.4 eV.

The resulting LF at $z = 6$ is shown in Fig. 3.2. The LF has a sharp cut-off at both faint ($M_{AB} > -10$) and bright ($M_{AB} < -18$) magnitudes, with a plateau and peak near $M_{AB} \sim -15$. There is a sizeable contribution from faint stellar sources, as well as a population of sources that are brighter than most BHs, except for the very brightest BHs at $M_{AB} \sim -23$. The LF from the simulation matches observational constraints well (Bouwens et al., 2015a), showing a similar distribution of brighter stars and reproducing the sharply rising behaviour of a Schechter LF. At the fainter magnitudes, where no observational constraint is available, the LF decreases, a feature shown to be related to a lower SFR and stellar content in smaller mass halos (see e.g. O'Shea et al. 2015). A reduction in SFR in low-mass galaxies due to photoionization feedback has also been found in fully coupled reionization simulations (Dawoodbhoy et al., 2018).

Nuclear Black Holes (BHs): In MBII each BHs is modelled as a collisionless sink particle of mass $5 \times 10^5\, h^{-1}\, M_\odot$ seeded within newly formed halos with mass above $5 \times 10^{10}\, h^{-1}\, M_\odot$. The BH seed grows by accreting surrounding gas or by merging with other BHs. The

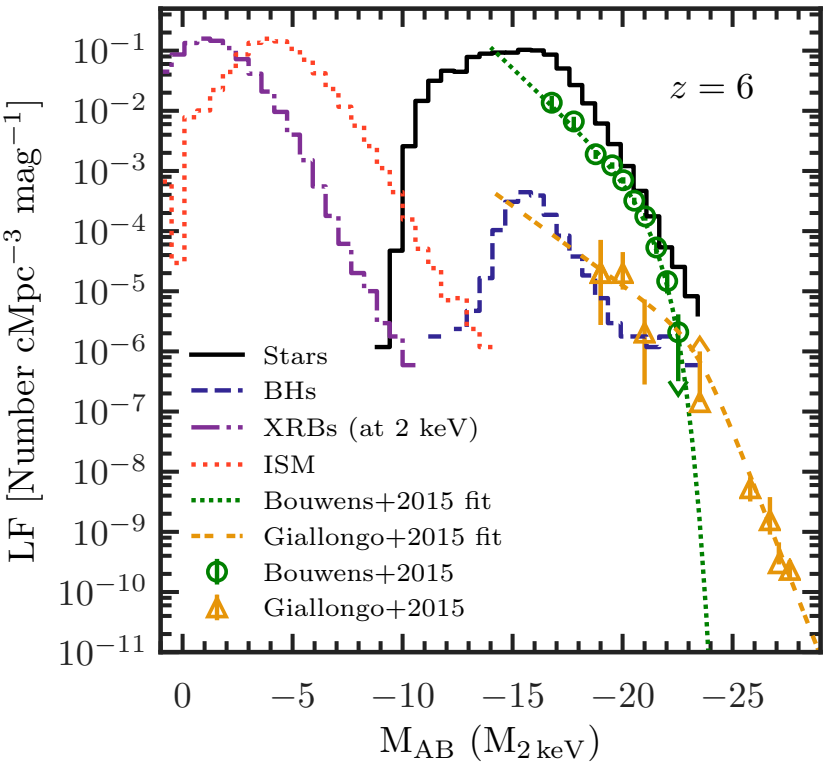

Figure 3.2: Luminosity functions at AB magnitude 1450 Å at $z = 6$ for stars (solid black line), BHs (dashed blue), ISM (dotted red) and XRBs (at 2 keV/6.2 Å, dash dot-dotted purple). The green circles and yellow triangles are observational data for stars (Bouwens et al., 2015a) and BHs (Giallongo et al., 2015), respectively, while the dotted and double-dashed lines are corresponding fits.

accretion rate, $\dot{M}_i$, depends on the BH mass, its velocity relative to the surrounding gas and the density and sound speed of the gas. $\dot{M}_i$ is mildly super-Eddington. Each BH in MBII is assumed to radiate with a bolometric luminosity $L_i^{\rm BH} = \eta \dot{M}_i c^2$ (in erg s^{-1}) given by its accretion rate $\dot{M}_i$ (in g s^{-1}) scaled by an efficiency parameter η and the speed of light c (Shakura & Sunyaev, 1973). The efficiency parameter $\eta = 0.1$ is chosen to be consistent with MBII. For all BHs we adopt an averaged spectrum obtained from observations of 108,104 low-z QSOs (Krawczyk et al., 2013), which is a broken power law with a spectral index of 1 for $h_{\rm P}\nu > 200$ eV. We rescale it with the luminosity of each BH, which rarely exceeds the Eddington value. A notable exception is our brightest BH, found at $z = 7$, which has an Eddington ratio of 2.9. In general, as can be seen from Fig. 3.1, the volume average of the BH spectra are all brighter than those of other source types. At $z = 6$ we have a total of 2,745 BHs. In Fig. 3.2 the BH LF at $z = 6$ is shown, along with the observational data (and a fitting LF) from Giallongo et al. (2015). Although we are constrained by the box size and therefore do not sample the rarest, brightest BHs, we do

however have a population of faint BHs. Our LF matches the range of brighter magnitudes where observational constraints are available. In chapter 4 by means of a neural network applied to the BH population of the MBII simulation, we re-estimate the faint end of the LF.

X-ray Binary Systems (XRBs): As with the black holes, we separately obtain a spectrum and rescale it with the individual luminosities of the sources. We look up spectra (evolving only with redshift) and luminosities (scaling with the physical properties of the sources) from the libraries of Fragos et al. (2013a,b), using the updated version presented by Madau & Fragos (2017). In particular, the luminosity scales with properties such as total mass, metallicity and age of the stellar particles, as well as the star formation rate in the halo. Our models include contributions from high-mass XRBs, which trace star formation and are dominant at $z \gtrsim 2.5$, as well as low-mass XRBs, which trace the stellar mass and dominate at lower z. Their spectra are generally hard, peaking at keV-scales, as shown in Fig. 3.1. As their AB magnitude is negligibly small, we calculate their $z = 6$ LF at a wavelength of 6.2 Å, corresponding to photon energies of 2 keV, and include it in Fig. 3.2. The peak of the 2 keV LF of the XRBs is at a magnitude $\mathrm{M}_{2\,\mathrm{keV}} \sim 1$.

Diffuse thermal bremsstrahlung from shock heated ISM (ISM): Whereas XRBs are point sources of X-rays in galaxies, a diffuse X-ray component has also been observed (Mineo et al., 2012b). We model this as a redshift independent broken power-law spectrum which is flat until the characteristic thermal energy break at $k_{\mathrm{B}}T^{\mathrm{ISM}} = 240$ eV (observationally determined by Mineo et al. 2012b, where k_{B} is the Boltzmann's constant), and has a spectral index of 3 at higher energies. As the shock heating process is likely to be associated with supernovae (Meiksin et al., 2017), we scale the ISM luminosities of the galaxies with their star formation rate following Mineo et al. (2012b). As seen in Fig. 3.1, the spectra are fainter than the stellar ones below HeII ionizing energies, while they are considerably softer than those of XRBs. From Fig. 3.2 we note that the ISM is much fainter than stars and BHs, where the brightest interstellar gas corresponds to the faintest BHs at $\mathrm{M}_{\mathrm{AB}} \sim -13$.

As we combine scaling relations from the literature with the physical properties obtained from MBII, the only real free parameter in our simulations is the escape fraction of ionizing photons. While in chapter 2 we rescaled the emission of $h_{\mathrm{P}}\nu < 200$ eV photons by the constant factor $f_{\mathrm{esc}} = 15\%$, observational and theoretical investigations suggest that the averaged escape fraction of ionizing photons may have evolved with redshift along with the changing occurrence of environments that allow for their escape[4]. In this work, we examine the effect of varying the escape fraction following the models and Bayesian constraints from Planck presented by Price et al. (2016),

$$f_{\mathrm{esc}}(z) = f_{\mathrm{esc},z=8} \left(\frac{1+z}{9}\right)^{\beta}, \tag{3.1}$$

[4]See e.g. observations from Izotov et al. (2018), Vanzella et al. (2018) and Fletcher et al. (2019), in addition to suggestions of higher f_{esc} due to turbulence by e.g. Safarzadeh & Scannapieco (2016), Kakiichi & Gronke (2019) and Kimm et al. (2019), or of an environment-dependent f_{esc} as seen in simulations from e.g. Paardekooper et al. (2015), Trebitsch et al. (2017) and Katz et al. (2018). A strong dependence is expected also on the evolution of the chemical properties of galaxies and in particular on the presence of dust (see e.g. Chisholm et al. 2018; He et al. 2020)

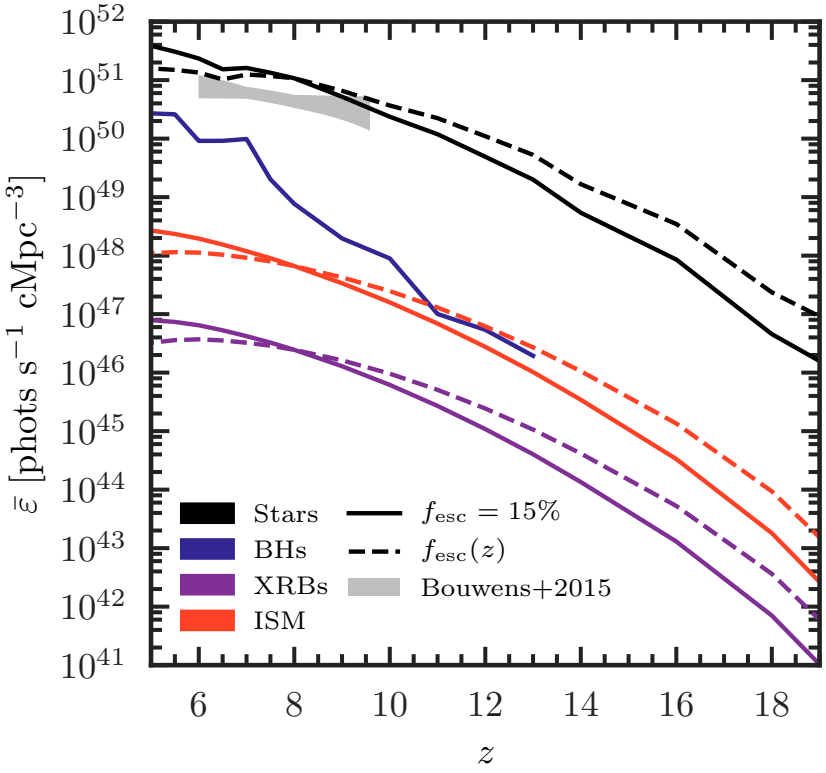

Figure 3.3: Redshift evolution of the comoving volume averaged ionizing emissivity per source type assuming either a constant UV escape fraction $f_{\rm esc} = 15\%$ (solid lines) or redshift-dependent $f_{\rm esc}(z)$ (dashed lines). From top to bottom the sets of curves correspond to: stars (black), BHs (blue), XRBs (purple) and ISM (red). The grey area corresponds to observationally constrained 95% CI for the evolution of the cosmic ionizing emissivity (Bouwens et al., 2015b).

where $f_{{\rm esc},z=8}$ is the escape fraction at $z = 8$. We choose $f_{{\rm esc},z=8} = 15\%$ and $\beta = 2.2$, where the latter is slightly lower than the Price et al. (2016) lower limit. Our evolving escape fraction is essentially 100% at high-z before decreasing to single-digit values at $z < 8$.

Our choices for the value of the escape fractions have also been guided to yield emissivities consistent with the Bouwens et al. (2015b) measurements, as can be seen in Fig. 3.3, where we show the evolution of the volume averaged ionizing emissivity from the various source types. The difference between a constant and varying escape fraction is now reflected on the emissivity, which is higher at $z > 8$ with $f_{\rm esc}(z)$. Stars dominate the budget at all z, as expected, independently from the adopted escape fraction. There is however an increasingly significant contribution from black holes at $z < 10$, as their volume averaged emissivity increases by three orders of magnitude from the appearance of the first BH at $z = 13$, and $z = 6$. The contribution from the ISM is similar to that of BHs between $z = 13$ and 11 assuming $f_{\rm esc}(z)$, otherwise it is sub-dominant. The volume averaged emissivity of

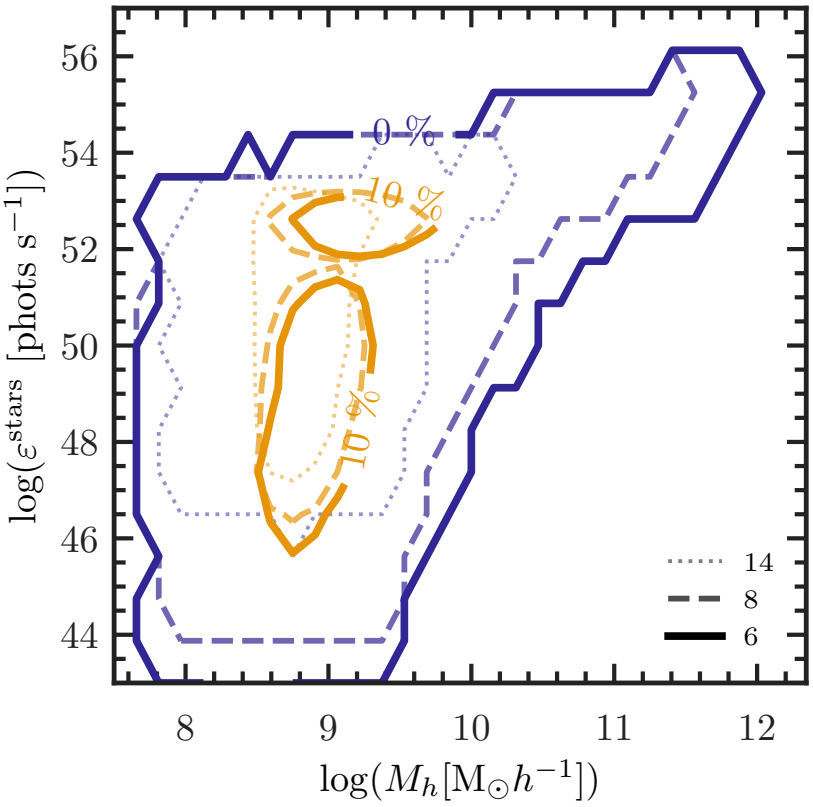

Figure 3.4: Ionizing emissivities of stars, $\varepsilon^{\rm stars}$, as a function of their halo masses, M_h, at $z = 14$ (thin dotted line), 8 (dashed line) and 6 (thick solid line). The contours are drawn at the 0th percentile (i.e. full distribution, blue lines) and 10th percentile (yellow lines).

the XRBs is fainter by yet another order of magnitude compared to that of the ISM.

To investigate which halos contribute most to the ionization budget, in Fig. 3.4 we plot the stellar emissivity (i.e. produced by all stellar particles contained within a halo), $\varepsilon^{\rm stars}$, as a function of the hosting halo mass, M_h. Halos are identified with a friends-of-friends procedure and have a minimum mass of $\sim 9 \times 10^6\, h^{-1}$ M$_\odot$. However, we have stars that have been associated (potentially unreliably) with even lower mass halos. We do not consider this to have significant impact on the reliability on the emissivity of the stars, as their properties are derived mainly from the star particles and not the halos. As $\varepsilon^{\rm stars}$ is assigned according to a number of physical properties, it has a strong degeneracy with M_h. The majority of the sources (90%) are found in the narrow range of halo masses $8.5 < \log M_h/({\rm M}_\odot h^{-1}) < 9.6$ at all redshifts considered here, but their emissivities differ by eight order of magnitudes, indicating the strong contribution from e.g. stellar ages and metallicities to the stellar emissivity. The most massive halos, with $\log M_h/({\rm M}_\odot h^{-1}) > 11$, are few in number, but yield exclusively high emissivities. These results are also suggestive that adopting the same escape fraction for all halos is an oversimplification.

As in chapter 2, we perform multifrequency 3D Monte Carlo RT with CRASH (Ciardi et al., 2001; Maselli et al., 2003, 2009; Graziani et al., 2013, 2018). We trace packets of photons that are emitted at the site of the sources. The emitted packets of photons are depleted and redshifted as they propagate (see chapter 1 for a more detailed description), and they are assumed to be lost once they exit the volume. This assumption does not impact the results presented here (see Appendix B.1 for a more extensive discussion).

3.3 Results

We first present some key qualitative findings that manifest themselves in all our results (subsection 3.3.1), and then turn to present detailed reionization histories, phase diagrams and thermal properties of the IGM in our simulations with various combinations of different source types and escape fractions (subsection 3.3.2). We finally study the results in light of various observational constraints (subsection 3.3.3).

3.3.1 Qualitative overview

In Fig. 3.5 we present a 3D rendering of the full $(100h^{-1}$ cMpc$)^3$ volume at $z = 6.9$, when the hydrogen reionization process nears completion and with helium reionization well underway[5]. It shows the HI and HeII ionization state of the IGM for a simulation including all the source types and with $f_{\rm esc} = 15\%$. The volume averaged ionization fraction is $\langle x_{\rm HII}\rangle = 0.85365$ for hydrogen, and $\langle x_{\rm HeIII}\rangle = 0.00731$ for HeII[6] The neutral IGM is perforated by ionized HII regions, not necessarily centered on the brightest sources. Instead, their extent illustrates the effect of the ionization provided by stars. Multiple galaxies contribute jointly to enlarge the HII fronts after individual HII bubbles merge. The morphology thus displays signatures of the percolation inherent to the HI reionization process (Furlanetto & Oh, 2016; Bag et al., 2019), a trait which is investigated in more detail in a companion paper (Busch et al., 2020). The HII regions host HeIII bubbles of various extent, most centered on BHs, indicating their responsibility in driving HeII reionization. The figure highlights our most luminous BH, which distinguishes itself through a large HeIII region. The spikes associated to it are mainly due to recently ionized channels of lower optical HeII depth, but there is also a component associated to numerical artefacts of a yet unconverged, non-equilibrium physical state (see also the physical properties of the environment of the high-z BH modelled by Kakiichi et al. 2017).

In Fig. 3.6 we show maps of the thermal and ionization state of the IGM at $z = 7.5$, centered around the brightest BH seen in Fig. 3.5. The BH resides in a 5.76×10^{11}M$_\odot$ halo, has an AB magnitude at 1450 Å $\mathrm{M_{AB}} = -21.9$, a mass $M_{\rm BH} = 1.85\times10^7M_\odot$ and an ionizing

[5]Note that the redshift displayed here is lower than those analysed in subsequent figures as in 3D more details can be discerned when the ionized fractions are higher. The presence of an ionized region at the location of a bright BH does not necessarily mean that the BH is active at all redshifts, but rather that the ionized region hasn't recombined yet.

[6]We reach a convergence at the 10^{-5} level. See chapter 2 for further details.

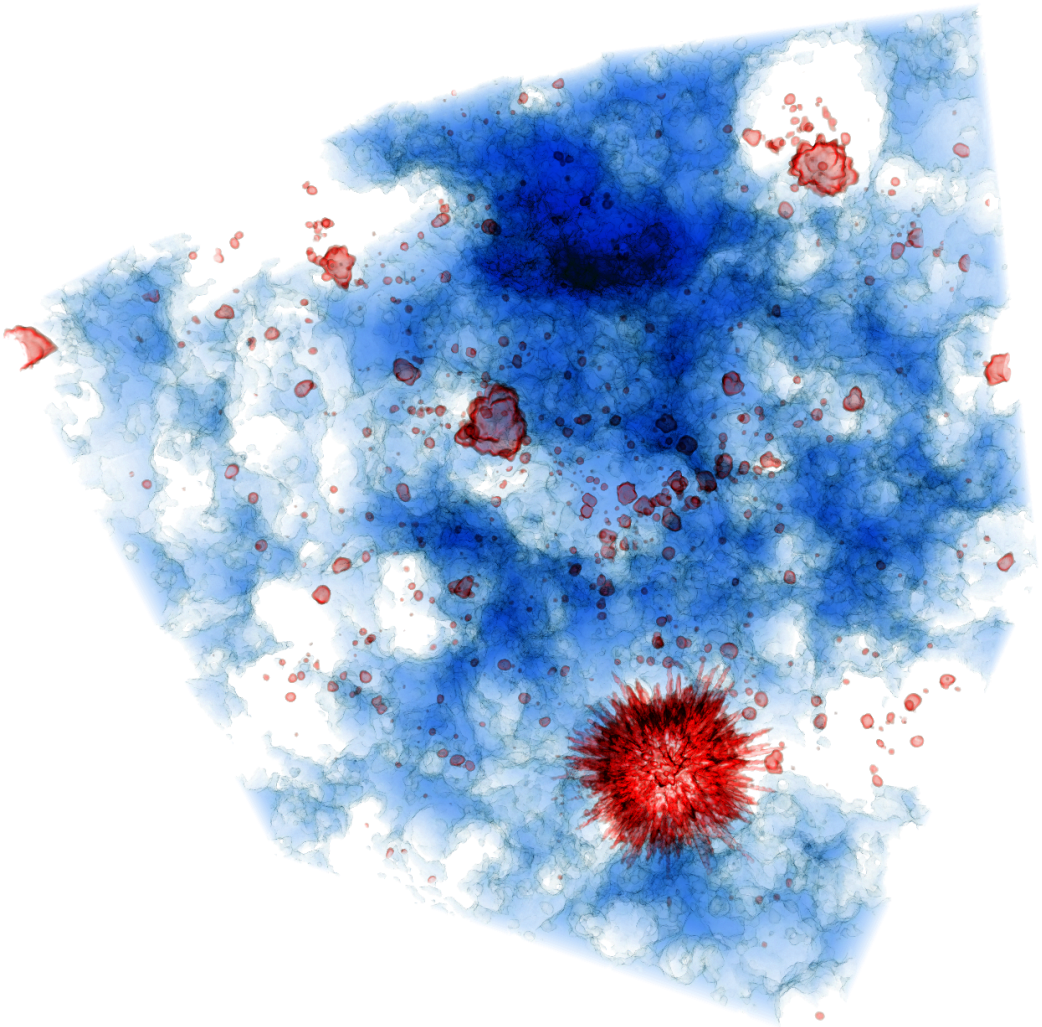

Figure 3.5: Volumetric rendering at $z = 6.9$ for a simulation with ionizing radiation from stars, BHs, XRBs and the ISM of galaxies, using a UV escape fraction $f_{\rm esc} = 15\%$. Each side of the volume is $100h^{-1}$ cMpc long. The blue regions show the distribution of the neutral HI and are drawn at $x_{\rm HI} > 0.1$, whereas red contours indicate $x_{\rm HeIII} > 0.1$, which we only find around BHs. The brightest BH is located in the lower right corner.

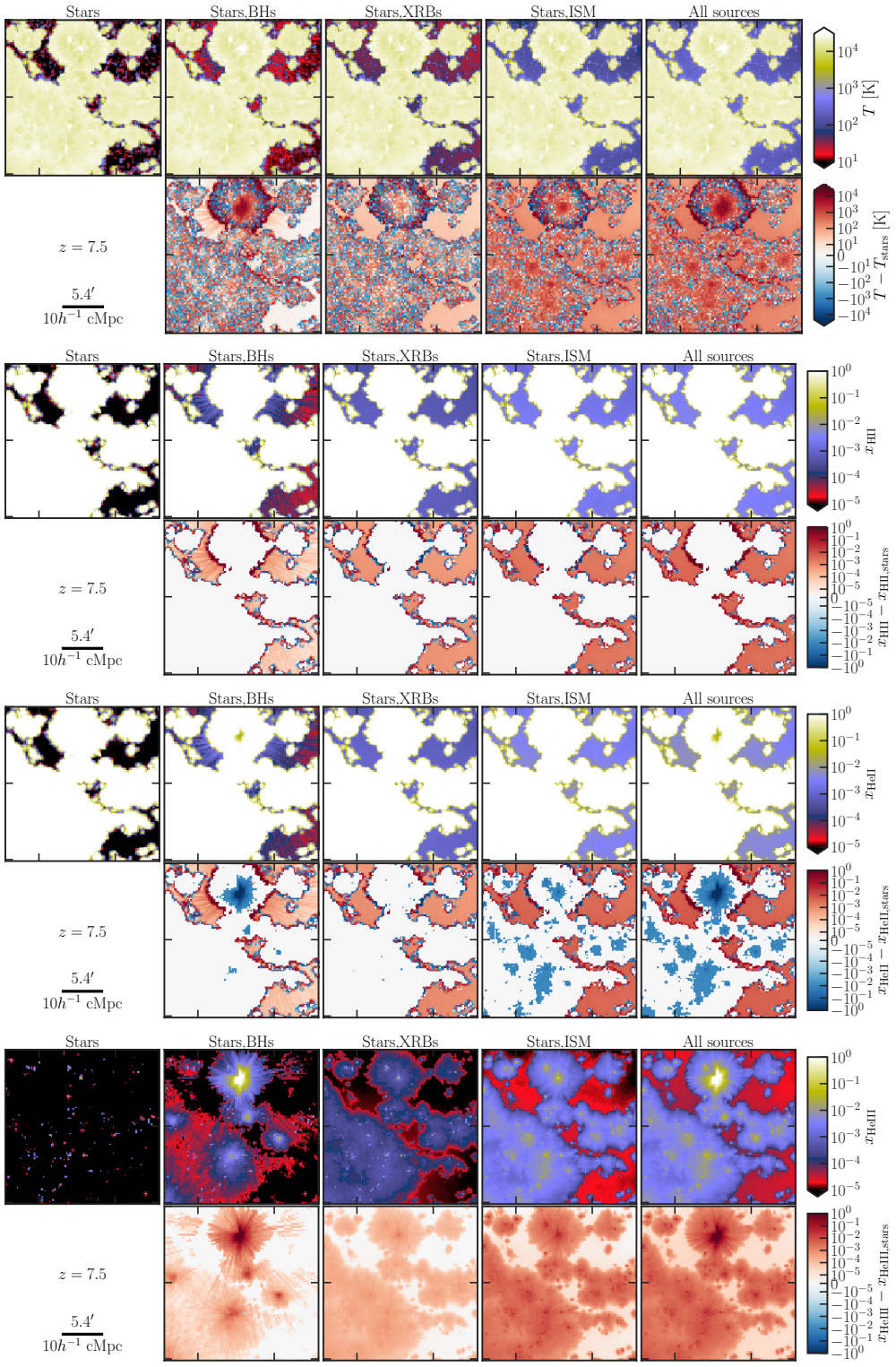

Figure 3.6: From top to bottom maps of the temperature, HII, HeII and HeIII fractions of the gas surrounding the location of a BH at $z = 7.5$ for different combinations of ionizing sources (left to right): stars, stars and BHs, stars and XRBs, stars and ISM, all sources. The lower sets of rows show the difference with respect to the stars alone simulations. The maps are $25h^{-1}$ cMpc wide. Note that alternating blue/red pixels at the edges of HII regions (and within the ionized regions in the temperature maps) are due to Monte Carlo noise and that only the red ring of extra ionization/heating produced by the bright quasar has a physical meaning. See text for more details.

emissivity $\varepsilon = 1.13 \times 10^{55}$ phots s^{-1}. This is not as bright or massive as ULAS J1120+0641 at $z = 7.54$ (which has $M_{AB} = -26.76 \pm 0.04$ and a mass $M_{BH} = 7.8^{+3.3}_{-1.9} \times 10^8 M_\odot$; Bañados et al. 2018), but it nevertheless belongs to the brightest end of the BH LF that can be sampled with a $100h^{-1}$ cMpc volume. The maps illustrate a few important concepts that will be further discussed in the following sections. The extent and temperature of $\sim 10^4$ K of the HII regions are determined by the stars. Additional and significant HII heating is only provided when HeII gets fully ionized by the BHs (see also Kakiichi et al. 2017 for further discussion), although partial ionization and heating around fully ionized regions is also provided by the shock heated ISM and, to a lesser extent because of the harder spectrum, the XRBs (see also Graziani et al. 2018). This will have a strong impact on 21 cm HI observations, which will be examined in Ma et al. in prep. Additional heating is also clearly visible within HII regions, in correspondence to a substantial presence of HeIII, i.e. in the immediate surroundings of the more energetic sources. The alternating blue/red pixels correspond to Monte Carlo noise, which is particularly evident in the temperature maps within the HII regions. This happens because it is statistically impossible for two cells with $x_{HII} = 1$ to have the same exact temperature in simulations with different source types.

Table 3.1: Thermal and ionization state of the IGM at $z = 6$ for different combination of source types and escape fractions. From left to right, the columns refer to the volume averaged fraction of HII, HeII, HeIII, HI, the volume averaged temperature, $\langle T \rangle$, and the temperature at mean density, T_0. The first and second value in each column refers to $f_{\rm esc} = 15\%$ and $f_{\rm esc}(z)$, respectively.

Source type	$\langle x_{\rm HII} \rangle$	$\langle x_{\rm HeII} \rangle$	$\langle x_{\rm HeIII} \rangle$	$\langle x_{\rm HI} \rangle$	$\langle T \rangle$ (K)	$\log(T_0)$/K
Stars	0.99998/0.99852	0.99971/0.99829	0.00019/0.00016	0.00002/0.00148	19,274/18,643	4.269/4.256
Stars, XRBs	0.99998/0.99854	0.99903/0.99762	0.00087/0.00085	0.00002/0.00146	19,374/18,746	4.271/4.258
Stars, ISM	0.99998/0.99869	0.98982/0.98889	0.01008/0.00972	0.00002/0.00131	19,975/19,350	4.284/4.272
Stars, BHs	0.99998/0.99923	0.98174/0.98110	0.01816/0.01804	0.00002/0.00077	19,524/18,915	4.275/4.263
All sources	0.99998/0.99935	0.97050/0.97035	0.02940/0.02892	0.00002/0.00065	20,351/19,751	4.293/4.281

3.3.2 Reionization and reheating history

In Fig. 3.7 we present the reionization and thermal histories for our five combinations of source types. The onset of HI and HeI reionization occurs at $z \sim 16$ for all models, and slightly earlier with an evolving escape fraction. With $f_{\rm esc}(z)$ the volume averaged ionization fractions remain higher until $z \sim 7$, when $f_{\rm esc}$ falls below 15% and the trend is reversed.

The main driver of the evolution of $\langle x_{\rm HII} \rangle$ and $\langle x_{\rm HeII} \rangle$ are the stars, with a non negligible contribution from the ISM and, at $z < 8$, the BHs (see Fig. 3.3). The ionization fractions differ most at $z \sim 7$, but they converge again to similar values towards the end of reionization, when $\langle x_{\rm HII} \rangle \sim 1$. We also observe that the HeII fraction below $z = 7$ with stars only is higher than with more energetic sources, as these are starting to produce an appreciable amount of HeIII, visibly depleting HeII.

The onset of HeII reionization is strongly dependent on the spectral hardness of the sources, occurring between $z \sim 11$ (when ISM or all sources are included) and $z \sim 9$ (with stars only). While XRBs increase $\langle x_{\rm HeII} \rangle$ by a factor of about five compared to stars only, including the contribution from the ISM or the BHs increases the difference to two order of magnitudes, the latter becoming the dominant source of HeII reionization at $z < 8$. With the exception of the initial and final stages of reionization, the evolution of the volume averaged temperature and ionization fractions are approximately exponential.

By $z \sim 6$ both HI and HeI reionization are concluded globally, with $\langle x_{\rm HII} \rangle \approx \langle x_{\rm HeII} \rangle \sim 1$ (see Table 3.1). However, the timing of reionization, $z_{\rm reion}$, varies with location, as it is clear from Fig. 3.8, where $z_{\rm reion}$ (defined as the redshift at which $x_{\rm HII}$ becomes larger than 0.9) is shown through a slice of our simulation with only stellar sources and $f_{\rm esc} = 15\%$ (see Fig. B.5 for a comparison with the other simulations). Reionization occurs earlier closer to the sources, supporting an inside-out scenario. Hydrogen in a small fraction of the IGM gets completely ionized at $z > 12$, including the one surrounding the massive BH discussed earlier, but reionization occurs at $z \sim 7$ for the majority of the IGM. Lowering the ionization threshold to $x_{\rm HII} \geq 0.1$ does not alter $z_{\rm reion}$ for most of the IGM, except that which is closest to the sources, where $z_{\rm reion}$ could increase from $z \sim 10$ to $z \sim 12$, as detailed in App. B.2.

Turning to the reheating history, its global onset, evolution and conclusion follows closely that of HI/HeI reionization. The temperature of the HII/HeII regions is determined mainly by stellar type sources, while the harder sources heat the IGM surrounding the fully ionized regions (as seen from the maps in Fig. 3.6). The volume averaged impact of these energetic sources is minor, but nevertheless visible, as the temperature is systematically higher in their presence (see Table 3.1). Similarly to what we found for HeIII, the difference with the stellar case only increases with decreasing redshift. Consistently with the trend observed in the ionization fractions, the temperatures are always higher with an evolving escape fraction. From the inset in Fig. 3.7, at $z < 6.5$ we observe a clear decline in temperature for all the simulations with $f_{\rm esc} = 15\%$, except when all sources are included, in which case it rises from $\langle T \rangle = 20,277$ K at $z = 6.5$, whereas this cooling is less evident with $f_{\rm esc}(z)$. The cooling is indicative of the transition towards an expected thermal

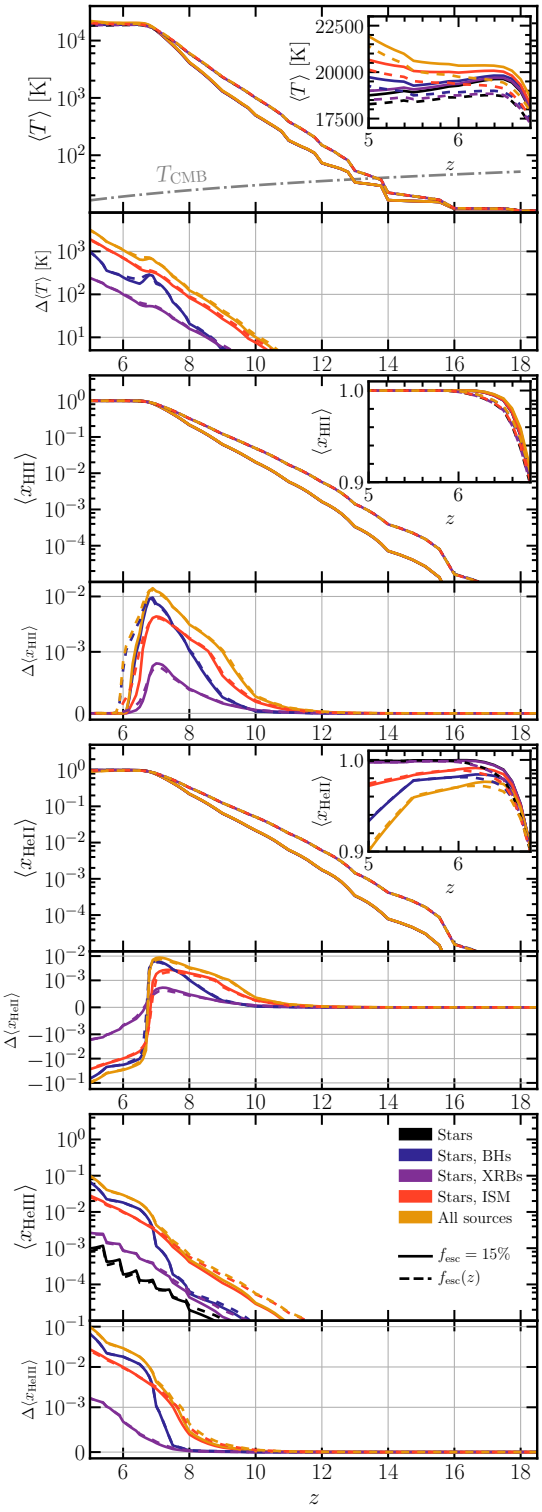

Figure 3.7: Evolution of the volume averaged (from top to bottom) temperature, fraction of HII, HeII and HeIII for a constant $f_{\rm esc} = 15\%$ (solid lines) and an evolving $f_{\rm esc}(z)$ (dashed). The line colour refers to the ionizing sources included: stars (black), stars and BHs (blue), stars and XRBs (purple), stars and ISM (red), and all sources (yellow). The lower panels show the differences with respect to simulations with only stars. The dotted-dashed grey line in the top panel is the CMB temperature.

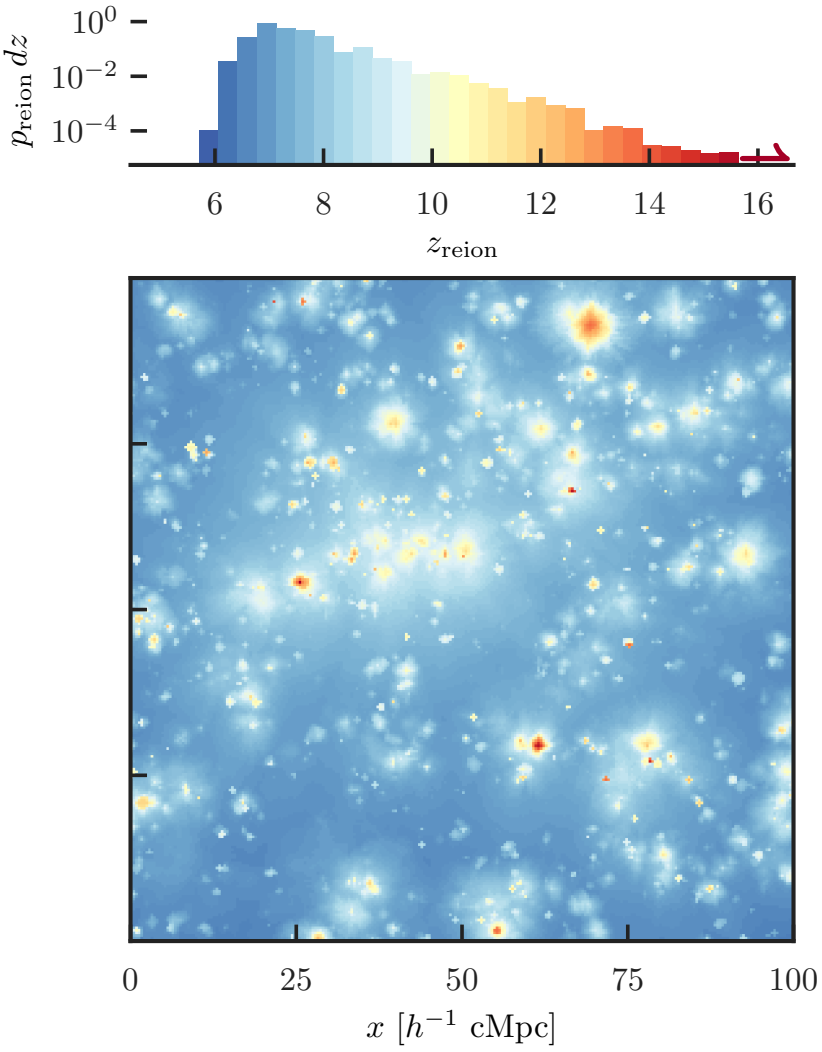

Figure 3.8: Slice of the simulation with stars only and $f_{\mathrm{esc}} = 15\%$ showing the reionization redshifts z_{reion}, colour coded according to the values in the histogram. The histogram indicates the distribution of z_{reion}, given as the normalized probability density $p_{\mathrm{reion}}\,dz$.

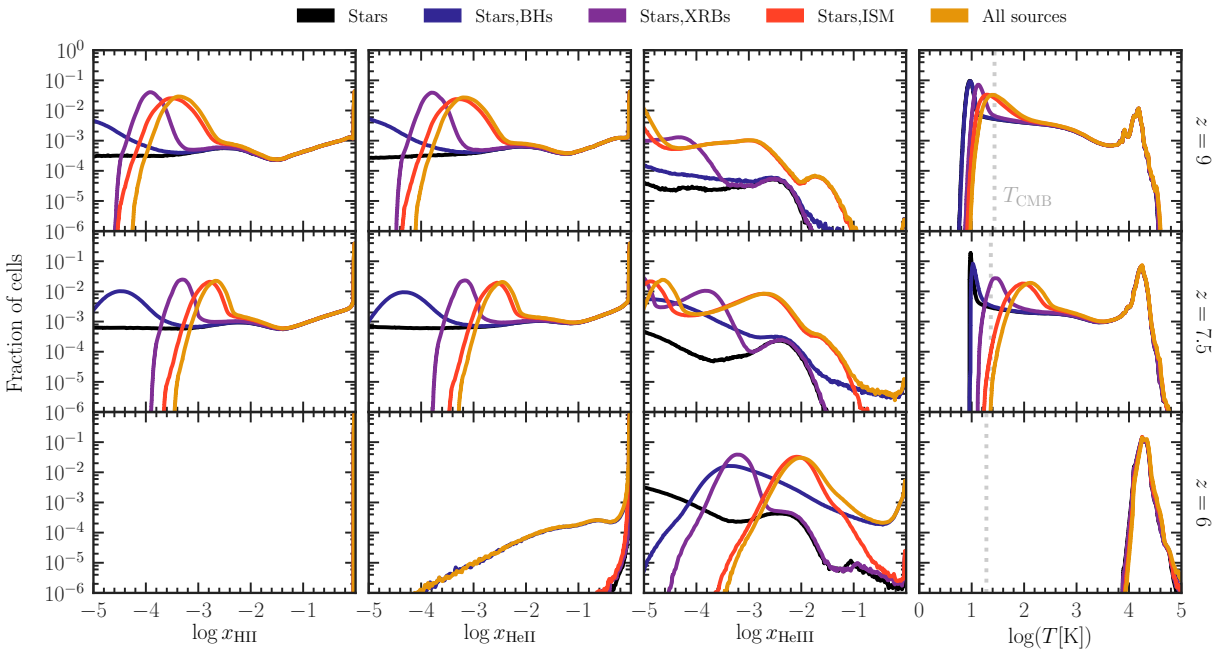

Figure 3.9: Distributions (from left to right) of the HII, HeII, HeIII fractions, and the temperature at (from top to bottom) $z = 9$, 7.5 and 6. The line colour indicates the type of sources considered: stars (black), stars and BHs (blue), stars and XRBs (purple), stars and ISM (red), and all source types (yellow). The vertical dashed grey line in the temperature panels indicate the CMB temperature, $T_{\rm CMB}$.

asymptote (Hui & Gnedin, 1997) which can be upset and postponed by HeII reionization, during which further heating is provided (McQuinn et al., 2009). At $z < 6$ we see the clear thermal signature of the latter in all our simulations except for those with stars alone, or stars and XRBs, both of which are inefficient HeII ionizers.

In Fig. 3.9 we plot the distributions of the various quantities characterising the physical state of the IGM at three different redshifts through reionization, $z = 9$, 7.5 and 6. We find pronounced bimodalities for all quantities except HeIII and while the IGM still has neutral hydrogen and helium. As previously discussed, full ionization of hydrogen and singly ionized helium, and heating of the gas to $T \gtrsim 10^4$ K, is caused by stellar type sources. The IGM that is not ionized, is coldest and most neutral with stellar type sources. Of the other source types in addition to stars, we find more ionization of HI and HeI and heating with BHs, which is further increased by the presence of XRBs. The most partial ionization and heating, though, is provided by the ISM. The distribution of HeIII is complex. With stars, $\log x_{\rm HeIII} < -1.4$ at $z = 9$ and 7.5, and higher ionization is barely seen at $z = 6$. The BHs do not provide large amounts of HeIII until $z = 6$, when they are responsible for a large amount of fully ionized HeIII. The XRBs produce HeIII in appreciable quantities, but at lower $x_{\rm HeIII}$ values than the stars, while the ISM provides copious amounts of HeII ionization resulting in two peaks which shift towards higher values with decreasing redshift,

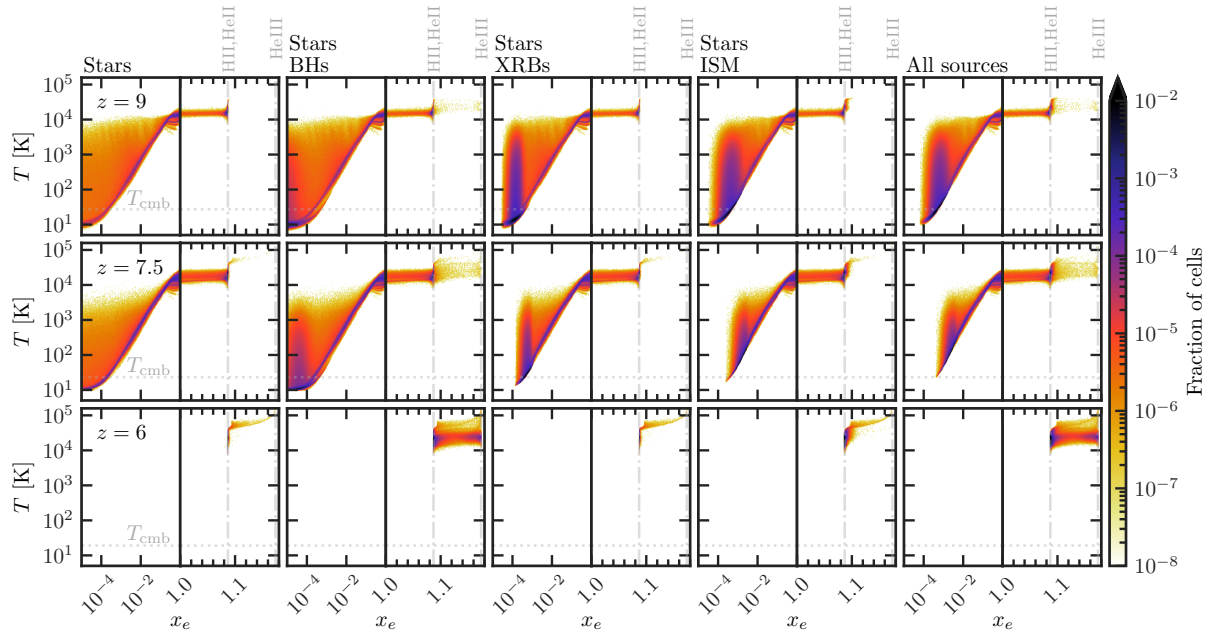

Figure 3.10: Distribution of the temperature, T, and the free electron fraction, x_e, for simulations with different combination of sources and $f_{\rm esc} = 15\%$ at $z = 9$ (upper row), $z = 7.5$ (middle row) and $z = 6$ (lower row). From left to right, the panels refer to stars only, stars and BHs, stars and XRBs, stars and the ISM, all sources together. The dotted horizontal lines indicate the CMB temperature, $T_{\rm CMB}$, while the dash-dotted vertical lines refer to full HI and HeI ionization ($x_e = 1.09$), and to full HI and HeIII ionization ($x_e = 1.17$), respectively.

but very little fully ionized helium.

We find gas with temperatures and ionization states in between the two peaks in the HII, HeII and T distributions. This gas may be a numerical artefact, as it is unlikely that such gas exists with stellar sources, as discussed by e.g. Ross et al. (2017) and Eide et al. (2018). Whenever the cell size is too large to resolve the sharp ionization front expected from stellar type sources, the cell containing the front appears partially ionized and warm, as in our simulations, while in reality part of the gas in the cell should be neutral and cold, and part fully ionized and hot. We discuss this issue more in detail in Ma et al. (2020) and Ma et al. (2020b, in prep).

In Fig. 3.10 we show the distribution of the volume with a given temperature T and free electron fraction

$$x_e \equiv \frac{n_e}{n_{\rm H}} = x_{\rm HII} + \frac{Y}{X}\left(x_{\rm HeII} + 2x_{\rm HeIII}\right), \tag{3.2}$$

where n_e and $n_{\rm H}$ are the number densities of free electrons and of hydrogen, while X and Y are the hydrogen and helium number fractions, respectively. As a reference, $x_e = 1.00$ when H is fully ionized and He is fully neutral, $x_e = 1.09$ when H is fully ionized and He is

singly ionized, and $x_e = 1.17$ when both H and He are fully ionized. We plot the relations at three different redshifts, $z = 9$, 7.5 and 6, showing distributions in which the gas is at various stages in transforming from cold ($T \sim 10$ K) and largely neutral ($x_e < 10^{-4}$) to hot ($T \sim 10^4$ K) and ionized ($x_e > 1$). At $z = 9$, all the simulations have a significant amount of gas with $x_e < 10^{-2}$ and temperatures ranging from 10 K to $\sim 10^3$ K, but still with significant amounts below $T_{\rm CMB}$. At $z = 7.5$, only the simulations with either stars alone or stars and BHs have gas with $T < T_{\rm CMB}$. At $z = 6$, all the gas in all the simulations has transitioned into an ionized state.

The highest temperatures (above $\sim 10^{4.5}$ K) are reached in regions where both H and He are ionized. Their abundance depends on the spectral shape of the ionizing photons and it is maximum in the presence of BHs. In fact, only BHs are able to fully ionize He (see third column of Fig. 3.9), although as a whole their contribution to ionization and heating is marginal because of their paucity. A slight amount of gas in this physical state is also present with stars only at $z = 7.5$ and 6 (see also the upper panel of Fig. 3.6), and corresponds to cells hosting sources[7].

The distribution of temperatures in the HII/HeII regions ($1 < x_e < 1.09$) is very similar in all simulations, ranging between ∼8,000 K and ∼20,000 K. The similarity confirms that stars are responsible for the physical state of the gas in such regions. The spread of temperatures observed in Fig. 3.10 is indicative of the time elapsed since the ionization of the gas. In fact, under purely adiabatic expansion and assuming no recombination, $T(z) \propto (1 + z)^2$. Furthermore, the cooling rate increases for temperature above $\sim 10^4$ K. The net effect is that gas that has recently been ionized is hotter than gas that has experienced ionization at an earlier time. This points to the gas temperature being a possible archaeological tracer of the reionization timing, supporting the work of e.g. Keating et al. (2018), who also found recently ionized regions to be hotter than earlier ionized ones.

As the temperature of the neutral and partially ionized IGM is of great importance to the cosmological HI 21 cm signal, we will discuss this further in the following. In Fig. 3.11 we plot the median temperature of gas with $x_{\rm HII} < 0.9$ against the volume averaged $\langle x_{\rm HII} \rangle$ for our different source types and escape fractions. The simulations reach $\langle x_{\rm HII} \rangle \sim 0.1$ at $z = 8$–10, when the CMB temperature ranges between 24.5 and 30 K. As expected, little or no heating of this largely neutral gas is observed with stars only. In spite of their harder photons' ability to penetrate deeper into the IGM, the contribution from BHs is negligible (a few degrees more than the stars) due to their paucity. The more numerous XRBs and ISM, individually fainter than BHs (see Fig. 3.2), are able to raise the temperature to $\sim$ 70 K and $\sim$ 220 K, respectively, as $\langle x_{\rm HII} \rangle$ approaches 0.9. An additional $\sim$ 50 K are gained when all sources are taken together. With stars alone, the temperature of the neutral IGM will always be below $T_{\rm CMB}$, whereas with BHs it will be patchy with temperatures above in the vicinity of BHs and below further away from these sources. As for the XRBs and ISM, we find $T_{\rm gas} \gg T_{\rm CMB}$ at all discernible $\langle x_{\rm HII} \rangle$values plotted in the figure, corresponding to redshifts $z \lesssim 10$. This behaviour has important implications for the 21cm signal, which will be presented in Ma et al. (2020b, in prep).

[7] A negligible fraction of these cells are shock heated.

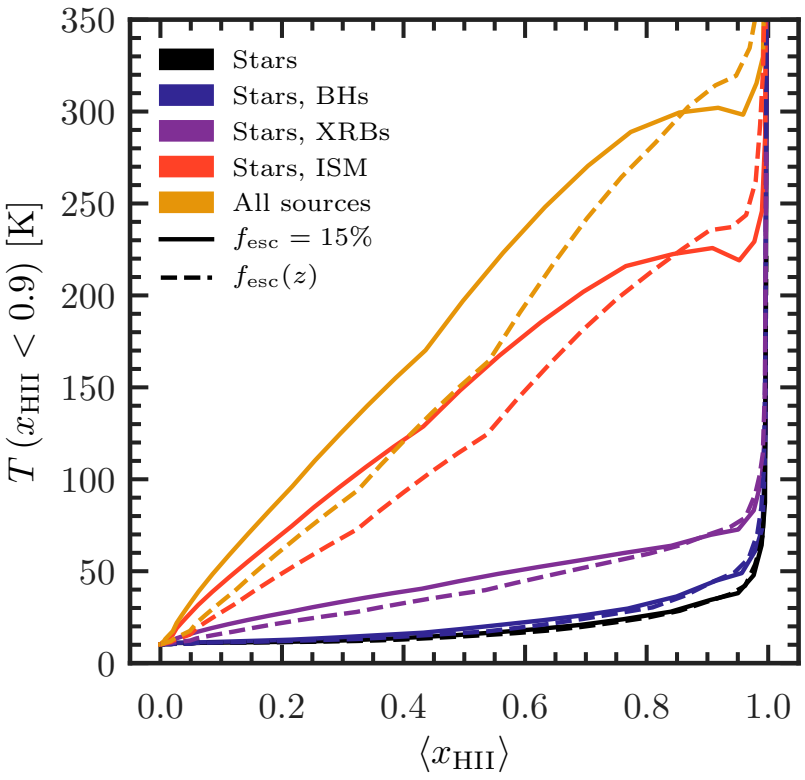

Figure 3.11: Median temperature of the gas with $x_{\rm HII} \leq 0.9$ as a function of $\langle x_{\rm HII} \rangle$. The solid and dashed lines refer to simulations with a constant $f_{\rm esc} = 15\%$ and a varying $f_{\rm esc}$, respectively. The line colour gives the combination of source types (black: stars; blue: stars, BHs; purple: stars, XRBs; red: stars, ISM; yellow: all sources).

Simulations with $f_{\rm esc} = 15\%$ have temperatures higher than those with $f_{\rm esc}(z)$ at a fixed $\langle x_{\rm HII} \rangle$ as long as $\langle x_{\rm HII} \rangle \lesssim 0.85$, while at higher ionization fractions the trend is reversed, as $f_{\rm esc}(z)$ at $z < 8$ reaches values lower than 15%, resulting in a hardening of the spectra in simulations with a varying escape fraction compared to those with $f_{\rm esc} = 15\%$, i.e. higher temperatures in the neutral parts of the IGM.

3.3.3 Observational constraints

In this section we will discuss our results in the context of available observational constraints.

Abundance of neutral hydrogen

In Fig. 3.12 we plot the evolution of $\langle x_{\rm HI} \rangle$ at $z < 12$ for simulations including all source types and the two different escape fractions (for reference, in Table 3.1 the values for all simulations are reported). Although the timing of full reionization is roughly the same

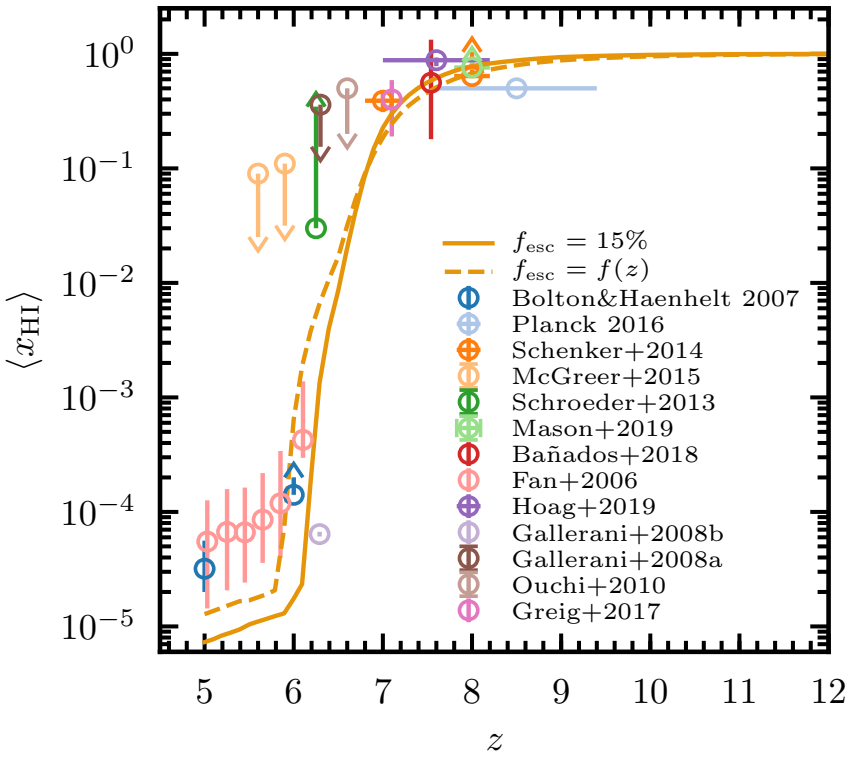

Figure 3.12: Evolution of the volume averaged neutral hydrogen fraction, $\langle x_{\rm HI} \rangle$. The lines refer to simulations with all sources and either a constant $f_{\rm esc} = 15\%$ (solid line) or an evolving escape fraction (dashed line). Observational constraints are also shown. These are discussed in the main text.

in both simulations, there are some differences in the residual neutral fraction towards its conclusion, with $\langle x_{\rm HI}\rangle = 6.5 \times 10^{-4}$ (1.8×10^{-5}) at $z = 6$ for $f_{\rm esc}(z)$ ($f_{\rm esc} = 15\%$). $\langle x_{\rm HI}\rangle$ remains approximately constant from $z = 12$ until $z = 9$, when it starts to decrease visibly. Consistently with what observed in Fig. 3.7, $\langle x_{\rm HI}\rangle$ is lower for simulations with $f_{\rm esc}(z)$ until $z \sim 7$, and then the trend is reversed, with a difference at $z = 6$ of about two orders of magnitude, as below $z = 6.5$ $f_{\rm esc}(z)$ has reached the single-digit percent level.

Our results are consistent with observational data at $z \gtrsim 7$, reproducing the $\langle x_{\rm HI}\rangle$ derived from absorption lines of the $z = 7.54$ QSO ULAS J1342+0928 by Bañados et al. (2018), the $\langle x_{\rm HI}\rangle$ statistically inferred from damping wing absorption in the spectra of ULAS J1120+0641 (Greig et al., 2017), the lower limit of Mason et al. (2019b) from the lack of Lyα emission in 53 faint LBGs at $z > 7.2$, the constraints from 68 LBG candidates of Hoag et al. (2019), and also the knee at $z \sim 7$ matches well the neutral fraction of Schenker et al. (2014) inferred from observed LBG visibilities.

At $z < 7$ our results are well within the upper limits from the dark pixel analysis of McGreer et al. (2015), from the LAE clustering by Ouchi et al. (2010), as well as the QSO sample analysis of Gallerani et al. (2008a). Nevertheless, at $z \sim 6$ we observe some tension. More specifically, the results with $f_{\rm esc} = 15\%$ are marginally consistent with the value of the neutral fraction derived along the line of sight to a GRB at $z = 6.29$ by Gallerani et al. (2008b), but both are at odds with the Fan et al. (2006) HI trough constraints at $z \leq 6$ and the QSO near-zones explored by Bolton & Haehnelt (2007), suggesting e.g. that toward the end of reionization a lower escape fraction would be more appropriate. The simulation with an evolving $f_{\rm esc}$ in fact seems to match better the observational constraints at $z \lesssim 6$. We should note that our estimate of the neutral fraction at these redshifts is likely an upper limit, as we do not resolve nor model as sub-grid physics the effect of small-scale Lyman-limit systems (LLS), which are instead crucial in the final stages of reionization as well as in the post-reionization IGM, as discussed e.g. by Madau (2017). The inclusion of LLS would also help to reconcile both our models with the QSO damping wing analysis by Schroeder et al. (2013), which indicates that a substantial part of the IGM must still be neutral at $z = 6$. Such neutral patches may also be required to explain the long troughs observed in front of QSOs (Chardin et al., 2017).

Finally, our timing of the reionization midpoint, $z_{50\%}$, when the volume average of the free electron fraction reaches $\langle x_e\rangle = 0.5$, is 7.5 for $f_{\rm esc} = 15\%$ (independent of source types) and 7.7 for $f_{\rm esc}(z)$, in both cases within the constraints of Planck Collaboration et al. (2018), where the middle 68th percentile is estimated to be $6.9 < z_{50\%} < 8.1$ with a tanh-parametrisation of $\langle x_e\rangle$.

IGM temperature

In Table 3.1 we also report the values of the temperature at the mean density, T_0. These are slightly higher than the $\log(T_0/{\rm K})$=4.21 $\pm$ 0.03 determined by Bolton et al. (2012) (the value is 3.9 $\pm$ 0.1 when correcting for the effect of HeII ionization within QSO near zones).

It should be noted that the IGM temperature is very sensitive to the spectral shape adopted for the sources. While our approach avoids the introduction of additional degrees

of freedom that comes with the assumption of a stellar spectrum (e.g. slope and shape), the results are strongly constrained by the hydrodynamic simulation[8].

In Fig. 3.13 we show the cumulative energy deposited as heat per unit mass in the simulations, $\langle u_0 \rangle$, calculated following Boera et al. (2019) and Nasir et al. (2016). This can be used as an independent constraint on reionization histories alongside the gas temperature. Independently from the source type, we observe a rapid increase in $\langle u_0 \rangle$, which roughly corresponds to the midpoint of reionization, $z_{50\%}$. This period of rapid heating is slightly shorter with a constant rather than a varying escape fraction, although in the latter case $\langle u_0 \rangle$ is ~ 0.2 eV $m_{\rm prot}^{-1}$ larger at $z \sim 8$, consistently with the evolution of the gas temperature. The rate of heating declines at $z < 6.6$ and it eventually becomes lower with $f_{\rm esc}(z)$ (see inset of the figure). At $z = 6$ all our simulations produce values of $\langle u_0 \rangle$ in the range $(5.58 - 5.93)$ eV $m_{\rm prot}^{-1}$, higher than those derived by Boera et al. (2019) from measurements of the Lyα flux power spectrum, with the exception of models with stars only and stars plus BHs with an evolving $f_{\rm esc}$. On the other hand, our results at $z = 6$ are all below the $\langle u_0 \rangle$ obtained by Boera et al. (2019) from the UV background (UVB) of Puchwein et al. (2019), while at $z < 6$, they seem to converge towards the Boera et al. (2019) and Puchwein et al. (2019) results at $z = 4.6$. Our $\langle u_0 \rangle$ histories show a more rapid increase in heating than Boera et al. (2019) find in their example models with the underlying Haardt & Madau (2012) UVB. However, the T_0 they derive in their models and from their measurements ($T_0 = 7,600$ K), as well as the $T_0 = 12,000$ K of Puchwein et al. (2019), are both lower than our values. The more rapid evolution of our $\langle u_0 \rangle$ and the higher T_0 can be understood as a consequence of the more rapid redshift evolution of the emissivity in our simulations, compared to those embedded in the UVB underlying the above estimates. Earlier ionization also gives ionized gas time to cool to a thermal asymptote. The majority of our IGM is reionized at later times (as seen in Fig. 3.8) to temperatures similar to those expected by Boera et al. (2019) (who estimate the gas temperature of recently reionized gas to be 20,000 K) and has hence not undergone a similar cooling.

Our wide range of temperatures in the post-reionized regions is consistent with the thermal properties of the ionized gas in the full radiation-hydrodynamics simulation investigated by D'Aloisio et al. (2019). As mentioned earlier, we find that the source spectral properties strongly affect also the thermal distribution, which is much more uniform in the presence of a power-law spectrum. This dependence on the spectral shape is not found in D'Aloisio et al. (2019), most probably because their spectra are cropped at 4 Ryd. We note that, when employing a full RT, D'Aloisio et al. (2019) find temperatures higher than when using uniform UVB models (such as of Puchwein et al. 2019), or RT simulations with low resolution or monofrequency spectra (e.g. Keating et al., 2018). This is consistent with test cases we have investigated. However, we do also find a wide range of temperatures on all scales in the post-ionization front zones (ranging from $\sim 18,000$ K to $\gtrsim 25,000$ K), highlighting that the simulated volume needs to be large enough for the results to be rep-

[8]We have tested that adopting a simple power-law spectrum results in lower volume averaged temperatures $\langle T \rangle$ and ionization fractions $\langle x_{\rm HII} \rangle$, whereas the neutral and partially ionized IGM is heated and ionized more.

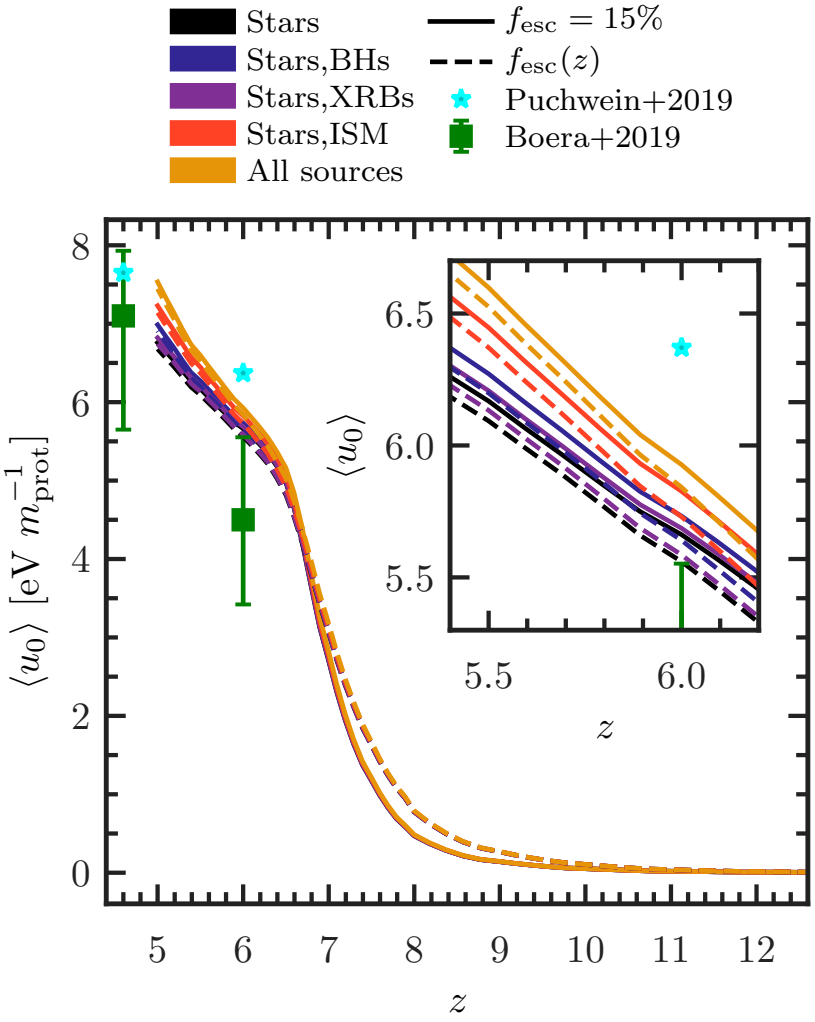

Figure 3.13: Cumulative heating per unit mass, $\langle u_0 \rangle$, of simulations with different source types (indicated by the line colour—black: stars; blue: stars, BHs; purple: stars, XRBs; red: stars, ISM; yellow: all sources) and escape fraction (indicated by the line type–solid: 15%; dashed: varying, $f_{\rm esc}(z)$). We also show observational constraints of Boera et al. (2019) (green squares) and Puchwein et al. (2019) (cyan stars).

resentative (Iliev et al., 2014). This does not rule out consistency between our findings and the lower T in works with smaller scales, such as Finlator et al. (2018). Our temperatures, though, are lower than those found by Ross et al. (2019) in their large scale RT simulations including QSOs and XRBs, where they predict $\langle T \rangle \sim 200$ K at $z = 14$, compared to the $\langle T \rangle \sim 10$ K of our models. However, our approaches differ in the way galaxies are populated with QSOs (as they use N-body simulations) and the choice of the spectra, so that a detailed comparison is not straightforward.

Finally, we note that the recent estimated constraints from the 21 cm observations of LOFAR at $z \approx 9.1$ (Ghara et al., 2020) find that, if the gas heating remains negligible, a mean ionization fraction of the IGM $\gtrsim 0.13$ is ruled out. These results align with the temperature inferences from PAPER-64 (Greig et al., 2016) and SARAS 2 (Singh et al., 2018), which rule out a cold reionization scenario. Fig. 3.11 suggests that the model including all source types is in closest agreement with the LOFAR constraints.

Thomson scattering optical depth

In Fig. 3.14 we show the optical depth τ due to electron scattering for our simulations together with the 68% CI optical depth $\tau = 0.054 \pm 0.007$ evaluated by the Planck Collaboration et al. (2018) and the slightly larger $\tau = 0.058 \pm 0.012$ of Planck Collaboration et al. (2016). The optical depth is calculated directly from the ionization fractions of our simulations at $z > 5$, while we assume $x_i(3 < z < 5) = x_i(z = 5)$ (where i= HII, HeII, HeIII) and that HeII reionization ends instantaneously at $z_{\rm HeII} = 3$, i.e. that $\langle x_{\rm HeIII} \rangle = 1$ for $z \leq 3$[9]. Similarly to what we have observed for the reionization history, stellar type sources determine the optical depth, producing $\tau = 0.0546^{+0.0003}_{-0.0003}$ ($0.0571^{+0.0003}_{-0.0003}$ with a varying $f_{\rm esc}$), compared to the marginally larger $\tau = 0.0549^{+0.0003}_{-0.0002}$ ($0.0575^{+0.0003}_{-0.0002}$) obtained for all sources. The simulations with a constant $f_{\rm esc}$ produce less ionization at high redshift, and consequently a lower τ than those with a varying $f_{\rm esc}$. We thus find that they are consistent with both the Planck Collaboration et al. (2018) and the Planck Collaboration et al. (2016) constraints.

3.4 Discussion and Conclusions

In chapter 2 we found that during the Cosmic Dawn there is a subtle interplay between various sources of ionizing photons. Their collective impact on the reionization process is not simply given by a sum of the single components. This is particularly relevant for a correct determination of the IGM temperature, but also for partially ionized H and He ionization. Here, we thus perform an analysis similar to that in chapter 2 concentrating on the Epoch of Reionization, at $z \lesssim 10$. This is done by modelling the reionization of hydrogen and helium by post-processing the outputs of the SPH simulation MassiveBlack-II (Khandai et al. 2015) with the 3D radiative transfer code `CRASH` (e.g. Ciardi et al. 2001;

[9]We evaluate τ also for $z_{\rm HeII} = 2.5$ and 3.5 and find that the results are negligibly affected by this choice.

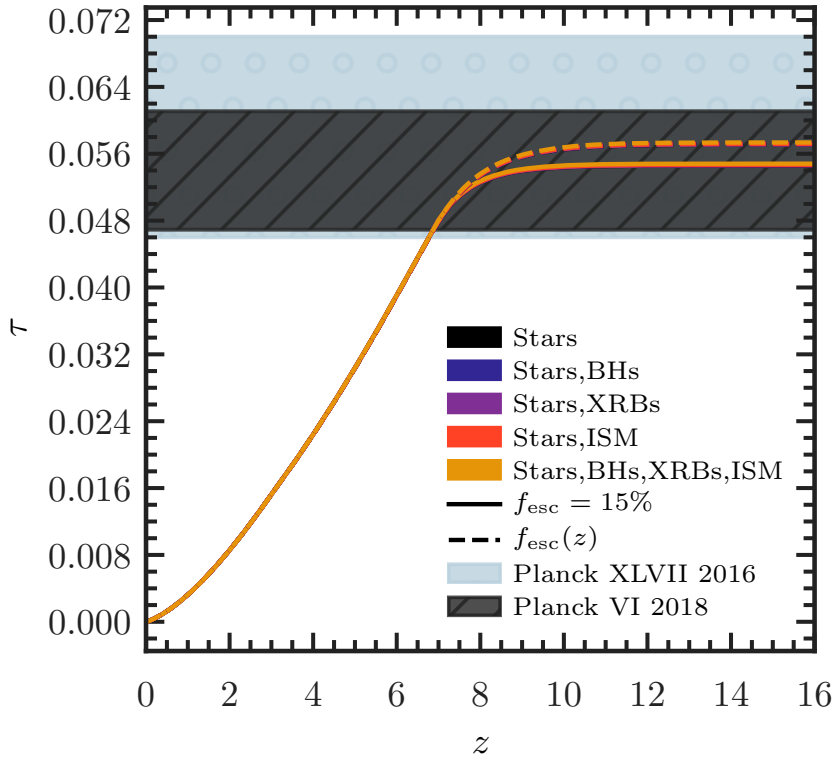

Figure 3.14: Evolution of the Thomson scattering optical depth, τ, for simulations with different source types (indicated by the line colour—black: stars; blue: stars, BHs; purple: stars, XRBs; red: stars, ISM; yellow: all sources) and escape fraction (indicated by the line type—solid: 15%; dashed: varying, $f_{\rm esc}(z)$). We note that lines referring to simulations with the same escape fraction basically overlap. The dark dashed and blue circled regions refer to the constraints from Planck Collaboration et al. (2018) and Planck Collaboration et al. (2016), respectively.

Maselli et al. 2003; Graziani et al. 2013, 2018), evaluating the physical properties of the IGM as determined by a variety of sources, namely stars, nuclear accreting black holes, X-ray binaries and shock heated ISM. The sources characteristics are derived directly from the properties of the stellar particles (such as stellar mass, age, metallicity for stars and XRBs), of the host galaxies (such as the star formation rate for the XRBs and the ISM) and of the BHs. Effectively, the only free parameters of the simulations is the escape fraction of UV photons.

Unlike most simulations in which the radiative transfer is coupled to the hydrodynamics (e.g. Gnedin 2014; O'Shea et al. 2015; Aubert et al. 2018; Rosdahl et al. 2018), our simulations give results on cosmological scales large enough to provide representative predictions for the progression of reionization (Iliev et al., 2014), although they are still not large enough to capture the effect of the largest and rarest QSOs and cannot match the large scales (and parameter space) explored by semi-numerical methods (e.g. Mesinger et al. 2011; Fialkov et al. 2014b; Hassan et al. 2017). Being less computationally expensive, though, our post-processing approach allows for a much higher sampling of the spectra of the sources in the radiative transfer, which is essential for a correct evaluation of the helium ionization and gas temperature in the presence of sources more energetic than stars. On the other hand, it fails to capture the feedback on structure formation from reionization itself. Our approach to post-processing differs in respect to those of other authors (e.g. Ross et al. 2017; Keating et al. 2018) as we use an SPH environment where the emissivities are not derived from dark matter halo properties, but rather from the properties of the sources as given by the hydrodynamic simulations (in essence, the only free parameter in our approach is the escape fraction of UV photons), in addition to having a very fine sampling of their spectrum.

Our results can be summarized as follows.

- Consistently with previous works and with chapter 2, full hydrogen reionization is driven by stars, which create rapidly growing and overlapping bubbles. The volume average HII fraction at $z \sim 6$ is in fact 0.99998 in all the scenarios investigated.

- Photons from more energetic sources (XRBs and ISM) propagate further into the IGM, inducing partial ionization and heating of the gas, in both its hydrogen and helium components, with values $x_{\rm HII} \sim x_{\rm HeII} \sim 10^{-4} - 10^{-2}$ and temperatures as high as $\sim 10^3$ K. Due to its soft spectrum, the ISM is particularly efficient. These results are consistent with what discussed in chapter 2 at higher redshift.

- While BHs do not have a strong impact on the average properties of the IGM due to their paucity, whenever active they dominate the local production of ionizing photons, increasing the extent of the HII regions. More importantly, they are the only sources capable of fully ionize helium (as it can be only partially ionized by XRBs and ISM), increasing the volume average HeIII fraction at $z \sim 6$ to ~ 0.02 compared to ~ 0.0002 with stars only, 0.0009 with XRBs and 0.01 with ISM.

- In the vicinity of energetic sources, where values $x_{\rm HeIII} \gtrsim 10^{-2}$ are reached, the gas temperature increases as much as 10^4 K compared to the case in which only radiation

from stars is considered. This confirms the importance of including the treatment of the helium component of the gas for a correct evaluation of the temperature. Similarly, because both the ionization level of helium and the value of the temperature are highly dependent on the spectral shape of the ionization field, it is equally crucial to sample the spectrum with a high accuracy.

- Among the models explored here, we find that simulations with an escape fraction that decreases with decreasing redshift reproduce more closely observational data, including the volume averaged emissivity. This also suggests that a constant value lower than the 15% adopted here could give a better fit. It should be noted, though, that the escape fraction is degenerate with properties such as the SFR and the stellar IMF. These are determined by the hydrodynamic simulations and the only free parameter thus remains $f_{\rm esc}$.

- Our models are consistent with constraints on the Thomson scattering optical depth from Planck, as well as on the volume average HI fraction down to $z \sim 6$, while a sub-grid treatment of the LLSs is necessary for a better modelling of the properties of the IGM at lower redshifts.

- Finally, our thermal histories are all dominated by the HII ionization of stars, reaching mean gas temperatures of $\langle T \rangle = 19,274$ K ($18,643$ K with an evolving escape fraction) with stars alone at $z = 6$. These temperatures only increase by at most $\sim 1,000$ K when including other source types. At $z < 6$ we see further heating from HeII ionization. Our temperatures are marginally larger than those deduced from observations, such as the temperature at mean density T_0 and the cumulative heating per unit mass $\langle u_0 \rangle$, and we conclude that this mainly is a feature of the redshift evolution of the ionizing emissivity. Important to the HI 21 cm line, the thermal state of the neutral or lowly ionized IGM is extremely sensitive to the presence and abundance of energetic sources. With only stars or additionally the BHs, we expect a 21 cm signal in absorption even at $z < 10$, whereas we expect a strong (XRBs) and stronger (ISM) signal in emission with the other source types. Observations (PAPER-64, SARAS 2, and LOFAR) favour our model with all source types combined.

Although clear differences in the IGM properties emerge depending on the source combinations, the observational data used in this chapter are not able to unambiguously discriminate between different source types and their relative contribution to the reionization process. Additional information is expected from the 21 cm line from neutral hydrogen (Ma et al. in prep), but, to maximize the extraction of information, it will be crucial to cross-correlate this signal with observations in different frequency bands such as of Lyα (e.g. Vrbanec et al., 2020) or OIII(Moriwaki et al., 2019) emitters, or of integrated quantities as X-ray (e.g. Ma et al., 2018b) and infrared (e.g. Fernandez et al., 2014) background, and the CMB (e.g. Jelić et al., 2010; Ma et al., 2018a).

Whereas redshifted OIIImay be observed with JWST, two other hyperfine transitions holds the potential to be probes of the high-z IGM (e.g. Furlanetto et al., 2006). The 3.46 cm transition of $^3He^+$ can be a unique signature of HeIII regions (Khullar et. al. in prep), which we have found to overwhelmingly exist only in the vicinity of BHs. Furthermore, this $^3He^+$ signal and the even more elusive DI 91.6 cm transition of deuterium may be observables less prone to the foreground contamination that occludes HI 21 cm signals.

To conclude, our large scale radiative transfer simulations of IGM reionization, anchored to detailed hydrodynamic simulations, give a new insight into the relative role of a variety of source types. The exceptional spectral resolution employed, as well as the inclusion of the helium component of the gas, also assure an accurate evaluation of the IGM temperature.

Chapter 4

The Black Holes during Reionization

4.1 Introduction

In the previous chapters, we simulated how our collection of source types could drive the Cosmic Dawn. While the stars were as predicted the sources mainly responsible for H reionization, the BHs became increasingly important towards lower redshifts. However, our population of BHs only stem from the most massive galaxies, as BH seeds of $M_{\rm BH,seed} = 5 \times 10^5 h^{-1}{\rm M}_\odot$ only were planted in galaxies once their halo mass exceeded $M_h \geq 5 \times 10^{10} h^{-1}{\rm M}_\odot$. This raises the question of the remainder of galaxies—should they host nuclear BHs as well, would these have a decisive impact on the Cosmic Dawn? The main challenge is however to reliably predict the population of BHs in ligher galaxies.

In this chapter we take a novel approach to model a high-z population of small BHs and study their impact on the EoR. This is done by training a neural network with the properties of the BHs and host galaxies modelled in the cosmological hydrodynamical simulation MassiveBlack-II (MBII, Khandai et al., 2015). The network is then used to mock the BH population (down to halo and BH masses lower than what was assumed and seeded in MBII) at redshifts relevant for the EoR. This work is published in Eide et al. (2020b).

The chapter is structured as follows: in section 4.2 we introduce the methods employed to develop the neural network; in section 4.3 we present our results in terms of BH and galactic properties, as well as the impact on the EoR; in section 4.4 we discuss some caveats and advantages of our new approach and give our conclusions.

4.2 Methods

In the following we will introduce the neural network adopted in our work.

4.2.1 Cosmological, Galactic and BH Properties

Here we present the 23 galactic and cosmological properties that we use as input to our neural network. From MBII (presented in the preceding chapters) we retrieve the stellar mass M_* (in $10^{10}h^{-1}\mathrm{M}_\odot$), the mean stellar metallicity Z, the star formation rate SFR (in $\mathrm{M}_\odot\,\mathrm{yr}^{-1}$), the mean stellar age τ (in yr), the dark matter halo mass M_h (in $10^{10}h^{-1}\mathrm{M}_\odot$), and the galactic gas mass $M_{\rm gas}$ (in $10^{10}h^{-1}\mathrm{M}_\odot$). We also derive some geometrical and kinematic properties of the galaxies by doing a principal component analysis of the velocities and positions of the gas and stellar particles (see e.g. VanderPlas et al., 2012). We find the galactic gas number density $n_{\rm gas}$ (in cm^{-3}), the mean velocity of the gas $\{\mu_r^{\rm gas}, \mu_\theta^{\rm gas}, \mu_\phi^{\rm gas}\}$ and of the stars $\{\mu_r^*, \mu_\theta^*, \mu_\phi^*\}$, and their respective velocity dispersion $\{\sigma_r^{\rm gas}, \sigma_\theta^{\rm gas}, \sigma_\phi^{\rm gas}\}$ and $\{\sigma_r^*, \sigma_\theta^* s, \sigma_\phi^*\}$, all in $\mathrm{km\,s}^{-1}$. For each galaxy we also have the stellar AB luminosity $L_{\rm AB}$ (in $\mathrm{erg\,s}^{-1}\,\mathrm{Hz}^{-1}$), and the stellar ionizing emissivity ε (phots s^{-1}), as calculated in Eide2018.

Additionally, we consider some cosmological properties at the site of each galaxy. Using the cosmic gas number density n, we calculate and grid onto 1024^3 regularly spaced cells the overdensity $\delta = n/\bar{n}$, where $\bar{n}$ is the volume averaged number density. As Di Matteo et al. (2017) found that the tidal field plays a central role in the growth of BHs, we follow their prescription to calculate and grid it. We evaluate the strain tensor in Fourier space, $\hat{S} = k^2\hat{\delta}/(k_i k_j)$ from the Fourier transform of the aforementioned gridded overdensity field δ (following Dalal et al., 2008), and find the tidal field as $T_{ij} = S_{ij} - \mathrm{Tr}\,S/3$. We calculate the eigenvalues of the tidal tensor, and retain the largest one, t_1. As we did for the overdensity, we read off t_1 from the grid at the site of the galaxy.

We additionally need to evaluate the accretion rate, luminosity and ionizing emissivity of the BHs. For this, we follow the approach taken in Eide2018 and Eide2020. In line with Shakura & Sunyaev (1973) and the feedback model employed in MBII, we write the bolometric luminosity as $L = \eta\dot{M}_{\rm BH}c^2$ (in $\mathrm{erg\ s}^{-1}$), where $\dot{M}_{\rm BH}$ is the BH accretion rate, $\eta = 0.1$ is an efficiency parameter and c is the speed of light. The ionizing emissivity $\epsilon_{\rm BH}$ (in phots s^{-1}) is derived by rescaling the integrated ionizing spectrum with the bolometric luminosity. The spectrum is determined observationally by Krawczyk et al. (2013) and it is essentially a broken power law at hydrogen-ionizing frequencies, with $L(\nu) \propto \nu^\alpha$ and $\alpha = -1$ for $h_{\rm P}\nu > 0.2\,\mathrm{keV}$, where ν is the frequency and $h_{\rm P}$ is the Planck constant. The integral of the ionizing spectrum gives the emissivity. From the rescaled spectrum we also derive the AB luminosity of the BHs, $L_{\rm AB,BH}$ (in $\mathrm{erg\ s}^{-1}\ \mathrm{Hz}^{-1}$).

4.2.2 The Neural Network

We now describe how we construct and train the neural network which ultimately is used to predict the BH masses, $M_{\rm BH}$, and accretion rates, $\dot{M}_{\rm BH}$. In essence, these are derived from the aforementioned 23 galactic properties, and the network is trained and validated on existing BHs at $z = 6$.

We use `TensorFlow`[1] with the `Keras`[2] interface to construct the neural network, while

[1] `https://www.tensorflow.org`
[2] `https://keras.io`

we employ `RMSProp` as optimiser. Starting from a single input layer $L(23)$ with the same number of units as we have learning parameters (23), we consecutively add layers with larger number of units to the network and test its accuracy after adding each new layer. We eventually arrive at a multilayered deep network where introduction of additional hidden layers lead to overfitting and modelling of the noise in the data because of the too many free parameters. We then introduce dropout layers, $D(r, v)$, which when enabled ($v = 1$) randomly remove a fraction r of the connections to the preceding layer, helping to increase the versatility of the network and to prevent overfitting (Hinton et al., 2012). The maximally connected network f can be described as one that takes an input vector $\mathbf{x}$ of our 23 learning (and prediction) parameters and forward feeds it through several hidden layers L and D before finally reaching an output layer which returns the predictions $\mathbf{y}^{\rm NN}$, where $y_0^{\rm NN} = M_{\rm BH}^{\rm NN}$ and $y_1^{\rm NN} = \dot{M}_{\rm BH}^{\rm NN}$. In its most complex form, it has the following structure,

$$\begin{aligned}\mathbf{y}^{\rm NN} &= f\left(\mathbf{x}; (LD(0.5,1))_3 L_3\right) = (LD)_3 L_3 \\ &= \mathbf{x} \rightarrow L(23) \\ &\quad \rightarrow D(0.5,1)L(92) \\ &\quad \rightarrow D(0.5,1)L(8) \\ &\quad \rightarrow D(0.5,1)L(16) \\ &\quad \rightarrow L(16) \\ &\quad \rightarrow L(2) \end{aligned} \tag{4.1}$$

where the arrows indicate that the outputs $\mathbf{a}^{l-1}$ of the layer $l-1$ are used to compute the activation of the units in the next hidden layer l, $\mathbf{a}^l = \texttt{ReLU}((1-r)^{-1}\mathrm{D}^l\mathrm{W}^l\mathbf{a}^{l-1} + \mathbf{b}^l)$. Here, W is the matrix of elements W_{ki}^l of the connection weights between unit i of layer $l-1$ and unit k of layer l, $D^l(r,1) = (1-r)^{-1}\mathrm{D}^l$ is a matrix where a fraction r of the connections are dropped, $\mathbf{b}$ is a bias, and `ReLU` is the activation function. The final layer has a linear activation function, i.e. $\mathbf{y} = \mathbf{a}^N = \mathrm{W}^N\mathbf{a}^{N-1} + \mathbf{b}^N$. In the following we omit 'r' from the notation as we always assume the standard vale $r = 0.5$ (Hinton et al., 2012).

In Fig. 4.1 we show how the combinations of the hidden layers affect the accuracy 1-MSE (mean-square error) of the M predictions, with $\mathrm{MSE} = M^{-1}\sum_{p=1}^{M}(y_{i,p}^{\rm NN} - y_{i,p})^2$ for $i = 0, 1$ and where y_i are the validation values. We test the network on the portion of the dataset that it has not been trained on. The figure shows combinations of $N = 1 \ldots 5$ hidden layers in addition to the output layer. A network where $N = 1$ only has two `ReLU` layers, $LD(v)_3 L$. We also show combinations of the dropout layers D. Networks where all the dropout layers are enabled are labeled as 'ddd' in the figure, whereas a combination such as e.g. $LD(1)(LD(0))_2 L_3 = LDL_4$ or $LD(1)D(0)_2 L = LDL$ is labeled 'dnn'.

The networks without any dropout layers, labeled 'nnn', usually have the smallest errors, but also the largest potential for overfitting the data. The networks with the smallest error (marked in the figure with a hatch) are L_4 for the BH mass, and L_5 for the BH accretion rate. The 'd**' networks, where a dropout is applied right after the input layer, generally present larger errors, particularly for the predicted BH accretion rates. Unsurprisingly, the predictions are better with dropout layers for the $N > 1$ layered

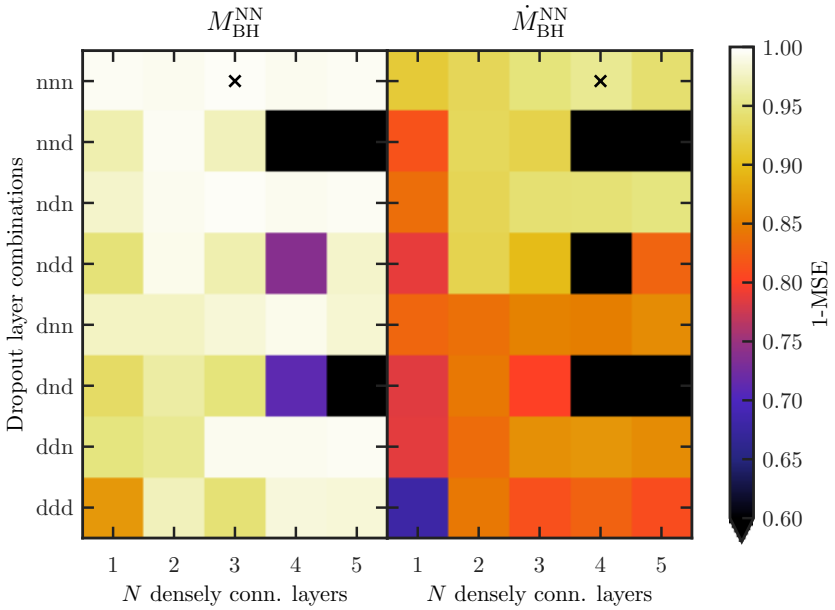

Figure 4.1: Impact of network composition on the accuracy 1-MSE in predicting BH masses $M_{\rm BH}^{\rm NN}$ (left panel), and accretion rates $\dot{M}_{\rm BH}^{\rm NN}$ (right panel). The x-axis shows the number of densely connected layers in the network, while the y-axis shows the combinations of dropout layers. The crosses indicate the networks with the highest accuracy.

networks. The 'ndn' class of networks particularly sets itself apart with consistently good predictions. We find that the LDL_2 network is the one that strikes the best balance between simplicity and predictive power, and is the one we apply in our work.

The training of all the networks was done at $z = 6$, where MBII has a sizable population of $2,734$ BHs. As the distribution of accretion rates is not uniform in our sample of galaxies hosting BHs, we whitened the input data before training. The whitening is done by duplicating galaxies with rare accretion rates, where the properties of the duplicates are added gaussian noise $\sim \mathcal{N}(0, 0.05\boldsymbol{\sigma})$ based on the variance $\boldsymbol{\sigma}$ of these properties. This extends the training sample and prevents the network from being biased towards only predicting the most common BH masses and accretion rates.

Furthermore, we did not train the networks on galaxies holding BHs with masses equal to those of the seeds, but rather restricted ourselves to $M_{\rm BH} > 1.1 M_{\rm BH,seed}$, as the BH properties just after seeding do not immediately reflect the properties of the host galaxy. This left 62% of the data available for training and verification. This also means that any prediction in the range $1 \leq M_{\rm BH}/M_{\rm BH,seed} \leq 1.1$ can be used to evaluate the predictive power of the network for masses that it has not been trained for.

We used 7/8 of the full sample of galaxies in the whitened set at $z = 6$ for training the network, before validating it on the remaining 1/8 of the set. By evaluating the MSE of the predictions of the network versus the validation data, we conclude that our ability to predict the BH masses and accretion rates with our chosen LDL_2 network happens with

an accuracy 1-MSE of 0.995 and 0.936, respectively.

4.3 Results

In this section we present our results in terms of galactic and BH properties, as well as the impact that our network predicted BH population has on the reionization process.

4.3.1 Relation between Galactic and Black Holes Properties

We now turn to examine if any of the 23 galactic properties plays a dominant role in predicting the BH masses and accretion rates. We do this by generating (i) a network which is the same as the original one except that now one parameter is removed, and (ii) a network using solely this parameter. In both cases we estimate the MSE on the predicted BH masses and accretion rates.

In Fig. 4.2 we show the networks' accuracy, 1-MSE, in predicting BH masses and accretion rates. A low $1-\mathrm{MSE}$ for the models plotted in red reflects a poorer performance of the network without the component under consideration. Conversely, a high $1-\mathrm{MSE}$ for the models in blue means that the predictive power of this single-parameter network is better.

We first note that the multi-parameter networks have a higher accuracy ($> 88\%$) compared to the single-parameter networks ($< 88\%$), and are hence performing better. The best single parameter for predicting the BH accretion rate and mass is M_*, for which the networks recover $M_{\rm BH}$ and $\dot{M}_{\rm BH}$ with an accuracy of 88% and 75%, respectively. As for the multi-parameter networks, they perform worst when removing M_*, yielding an accuracy of 99% and 89% for $M_{\rm BH}$ and $\dot{M}_{\rm BH}$, respectively.

While the relevance of M_* is clear both for single- and multi-parameter networks, this is not the case for the other parameters. The three next-most important parameters for the determination of $M_{\rm BH}$ are τ, ε and μ_ϕ^* for the multi-parameter network, whereas for the single-parameter network these are M, $M_{\rm gas}$ and $L_{\rm AB}$. Similarly, for $\dot{M}_{\rm BH}$ the most relevant quantities in the multi-parameter network are μ_ϕ^*, σ_r^* and $\mu_\phi^{\rm gas}$, indicating that the network captures the dependency between accretion and environmental kinematics; while $L_{\rm AB}$, SFR and $M_{\rm gas}$ yield the highest accuracies in the single-parameter networks. We recover the same order of importance for the single-parameter networks by calculating the correlation coefficient between $M_{\rm BH}$ or $\dot{M}_{\rm BH}$ and the parameter in question. Again, it should be noted that the predictions of these single-parameter networks are far less accurate ($\sim 60\%$) than the multi-parameter networks ($> 89\%$). As a further test, we create a network with only M_*, $L_{\rm AB}$, $M_{\rm gas}$ and M_h as input parameters. It recovers $M_{\rm BH}$ and $\dot{M}_{\rm BH}$ with an accuracy of 97% and 85%, respectively. This highlights our need for the full network's complexity if the goal is to recover the accretion rates as precisely as possible.

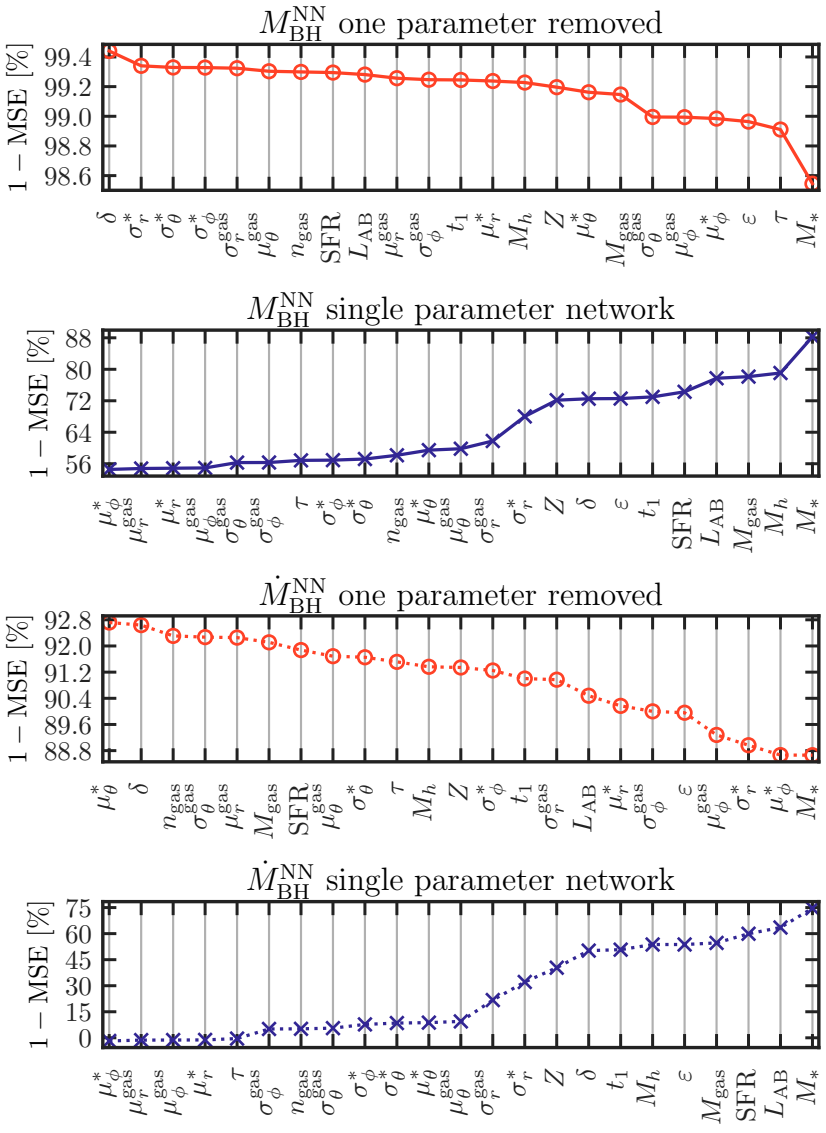

Figure 4.2: Network accuracy 1-MSE in predicting BH masses $M_{\rm BH}^{\rm NN}$ (solid lines), and accretion rates $\dot{M}_{\rm BH}^{\rm NN}$ (dotted lines). The red circled lines refer to the accuracy of the network in the absence of the parameter indicated in the x-axis, while the blue crossed lines refer to networks with only the removed parameter.

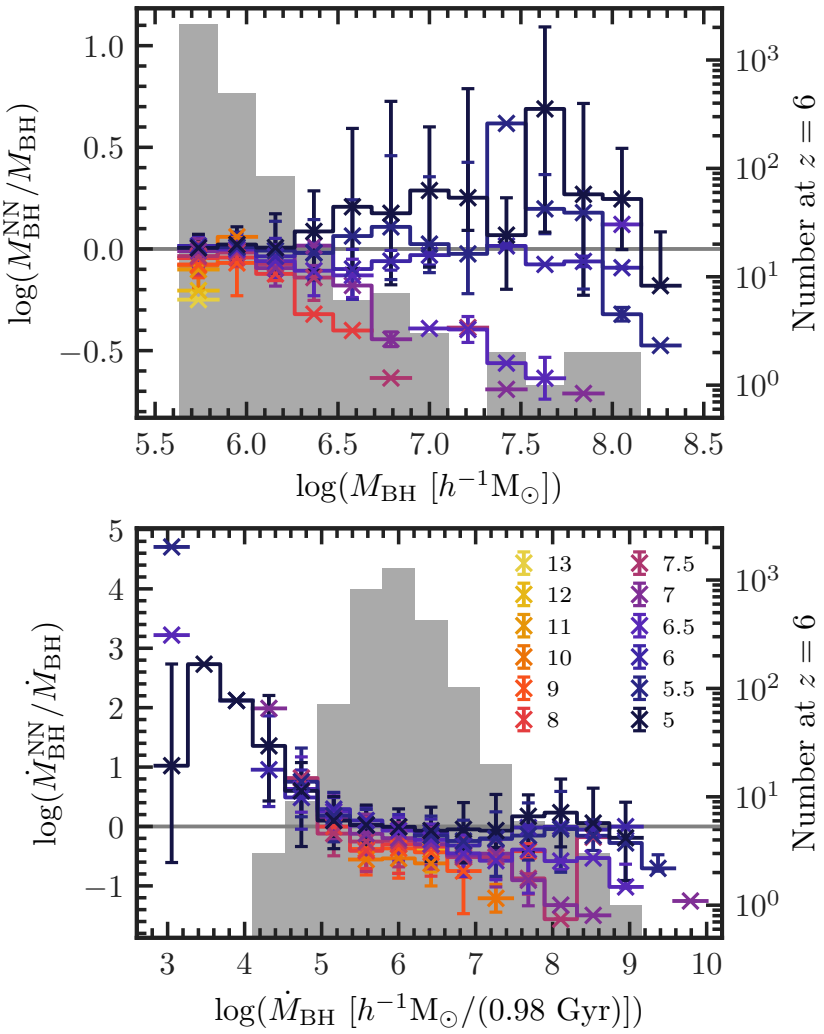

Figure 4.3: Deviation between the predictions of the neural network trained at $z = 6$ and the actual values, for galaxies harbouring BHs between $z = 13$ and $z = 5$, indicated by the line colour. The crosses, lower and upper limits give the median, 16th and 84th percentiles, respectively, of the associated bins. We show histograms of the masses and accretion rates of the MBII BHs at $z = 6$ in the background. The vertical dashed line indicates the lower training limit on the mass, $1.1M_{\rm BH,seed}$. *Upper panel:* predicted, $M_{\rm BH}^{\rm NN}$, against true, $M_{\rm BH}$, BH mass. *Lower panel:* predicted, $\dot{M}_{\rm BH}^{\rm NN}$, against true, $\dot{M}_{\rm BH}$, BH accretion rate.

4.3.2 Populating Galaxies with Black Holes

Our network was generated from $z = 6$ BHs and their host galaxies. We now turn to examine how it performs at other redshifts, for galaxies it has not been trained for. This can also reveal any redshift evolution in BH properties.

To do so, we use the network to seed BH-hosting galaxies at various z with BHs with predicted masses $M_{\rm BH}^{\rm NN}$ and accretion rates $\dot{M}_{\rm BH}^{\rm NN}$, and compare them to the $M_{\rm BH}$ and $\dot{M}_{\rm BH}$ directly obtained from MBII. We show the deviation between the generated and true values at various redshifts in Fig. 4.3. In the $M_{\rm BH,seed} < M_{\rm BH} < 10^{6.5} h^{-1}{\rm M}_\odot$ mass range, the deviations vary from 20% larger to 30% smaller, with the largest ones at $M_{\rm BH} > 10^7 h^{-1}{\rm M}_\odot$, where we have a poorer statistic of the training set. At $M_{\rm BH} < 1.1 M_{\rm BH,seed}$ where the network has not been trained for (this mass limit is indicated by a vertical dashed line in the figure), the predictions are 2–4% larger than the true values at $z = 6$, while at $z = 9$ they are $\sim 15\%$ lower. This indicates that our network is very powerful in predicting masses it has not been trained for. The predicted accretion rates deviates from being between $\sim 30\%$ larger to $\sim 60\%$ smaller for $10^5 < \dot{M}_{\rm BH}/(h^{-1}{\rm M}_\odot/(0.98\,{\rm Gyr})) < 10^{7.5}$. Also in this case, the predictions are best within the most common range of accretion rates. At the high mass and accretion rate end, the network underpredicts the true values at $z > 6$ and overpredicts them at $z < 6$. This indicates that the BH formation efficiency declines with decreasing z. We also see this effect within the central mass and accretion ranges, albeit in a much more moderate fashion—e.g. at $z = 9$ ($z = 5$), $M_{\rm BH} = 10^6 h^{-1}{\rm M}_\odot$ BHs are on average predicted to be 16% less (4% more) massive, and the accretion rate of $\dot{M}_{\rm BH} = 10^6 h^{-1}{\rm M}_\odot/(0.98\,{\rm Gyr})$ BHs is predicted to be $\sim 50\%$ ($\sim 4\%$) lower—indicating that this is not merely an effect caused by lacking statistics of our training set.

Next, we populate all galaxies in the range $z = 5 - 18$ with a BH using the neural network including all the 23 physical properties described in sections 4.2.1 and 4.2.2. We thereby create a much larger population of BHs than is present in MBII. In Fig. 4.4 we show the resulting mass function at various redshifts. While at $z = 18$ the BH population is limited to the range $10^4 < M_{\rm BH}/(h^{-1}{\rm M}_\odot) < 10^5$, the peak of the mass function shifts towards higher values with decreasing redshift, and by $z \sim 6$ we have BHs with masses as high as $\sim 10^8 h^{-1}$ M$_\odot$. The smallest BHs have $M_{\rm BH} \sim 10^{3.6} h^{-1}$ M$_\odot$ at all times. This is more than a magnitude lower than the seed mass $M_{\rm BH,seed}$ of MBII, and reflects that the predictions of the network are not restricted by the mass range it was trained on. Note that the generated BH mass function is not dissimilar to those for a range of physical BH seed models at $z = 15 - 18$ (e.g.; Volonteri et al., 2008). Our generated BH population appears to exploit reasonably well the actual resolution of the simulation, introducing BHs at smaller masses and earlier time when they are indeed expected to form. We note here that seeding halos of mass smaller than the one used in the MBII prescription is not a mere extrapolation, but is made possible by the fact that, even if the mass falls outside of the range used for the training, all the other 22 properties are not restricted by any limit. Hence the robustness of our procedure.

In Fig. 4.5 we show the UV luminosity function (LF) of the BHs at $z = 6$, and compare it to the LF of the MBII BHs, as well as to the observationally determined LFs of Giallongo

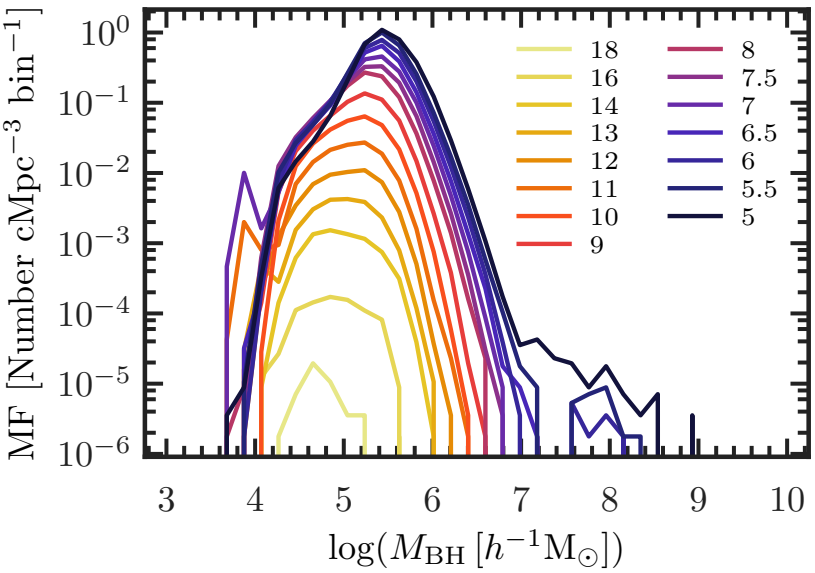

Figure 4.4: Mass function of BHs seeded with our neural network in every galaxy. The line color indicates different redshifts, in the range $z = 5 - 18$.

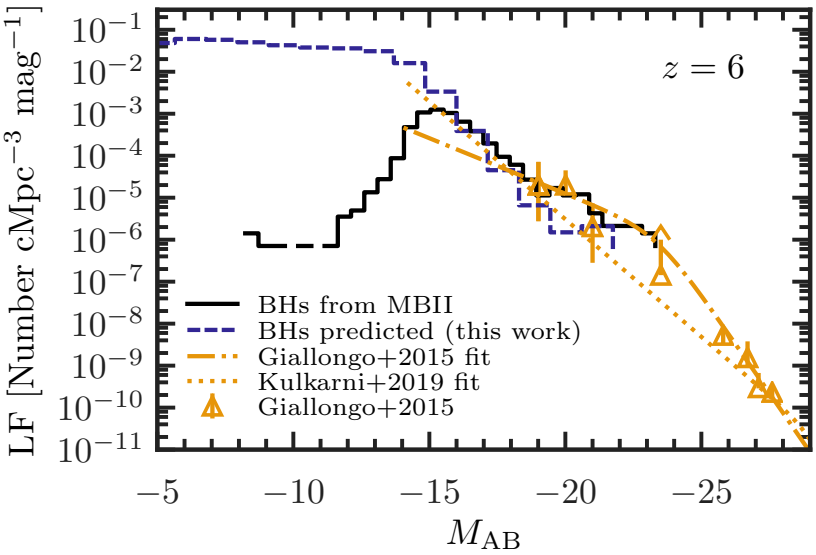

Figure 4.5: Luminosity functions at $z = 6$. The lines refer to the LF of the BH population from MBII (solid black line), the LF of the BHs seeded with our neural network (dashed blue line), and observational constraints from Giallongo et al. (2015) (data: yellow triangles, fit: dashed-dotted line) and Kulkarni et al. (2019) (fit: yellow dotted line) .

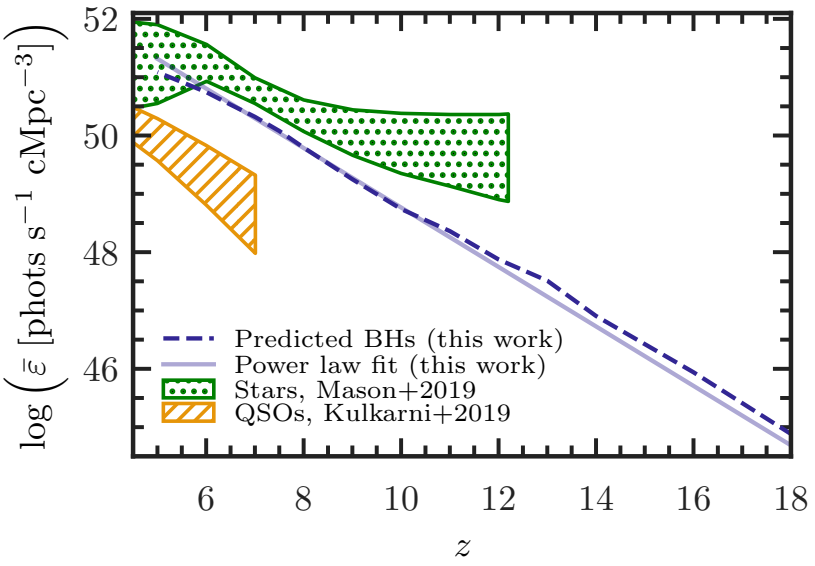

Figure 4.6: Volume averaged hydrogen ionizing emissivity $\bar{\varepsilon}$ (in phots s^{-1} cMpc^{-1}) from our predicted BHs (dashed blue line), with a power law fit (solid pale blue line). The hatched yellow area refers to the constraints derived by Mason et al. (2019a) based on observed bright QSOs compiled by Kulkarni et al. (2019). The upper (lower) limit includes QSOs with $M_{\mathrm{AB}} < -14$ (-21). The dotted green areas are inferences of the stellar contributions from observational constraints by Mason et al. (2019a).

et al. (2015) and Kulkarni et al. (2019). As our network slightly underpredicts the highest accretion rates, we have a small deficit of bright BHs compared to both the MBII-seeded BHs and the Giallongo et al. (2015) observations. It should be noted, though, that the bright end of the observed LF may be overestimated (Parsa et al., 2018), so that our conservative result might be more realistic. This is further corroborated by the recent compilation of Kulkarni et al. (2019), based on 66 QSOs at $5.5 < z < 6.5$, as our predicted LF matches their observations at all M_{AB}. The agreement of our LF with the original one from the MBII and the Giallongo et al. (2015) LF is extremely good in the range $-17 < M_{\mathrm{AB}} < -15$. Our network also predicts a substantial population of faint BHs which are not present in MBII, and yields a LF with a knee at $M_{\mathrm{AB}} = -15$ and no turnover at least down to $M_{\mathrm{AB}} = -5$.

In Fig. 4.6 we plot the comoving volume averaged emissivity, $\bar{\varepsilon}$, in comparison to values inferred from observations. The predicted emissivity increases exponentially from $z = 18$, when $\bar{\varepsilon} = 7.6 \times 10^{41}$ phots s^{-1} cMpc^{-1}, to $z = 5$, where $\bar{\varepsilon} = 1.2 \times 10^{52}$ phots s^{-1} cMpc^{-1}. This evolution can be parametrized as a power law,

$$\log \bar{\varepsilon}(z) = -0.5097z + 53.86, \tag{4.2}$$

using a least-square fit to the predictions. We find that the predicted emissivity is much higher than that inferred by Mason et al. (2019a) based on the Kulkarni et al. (2019)

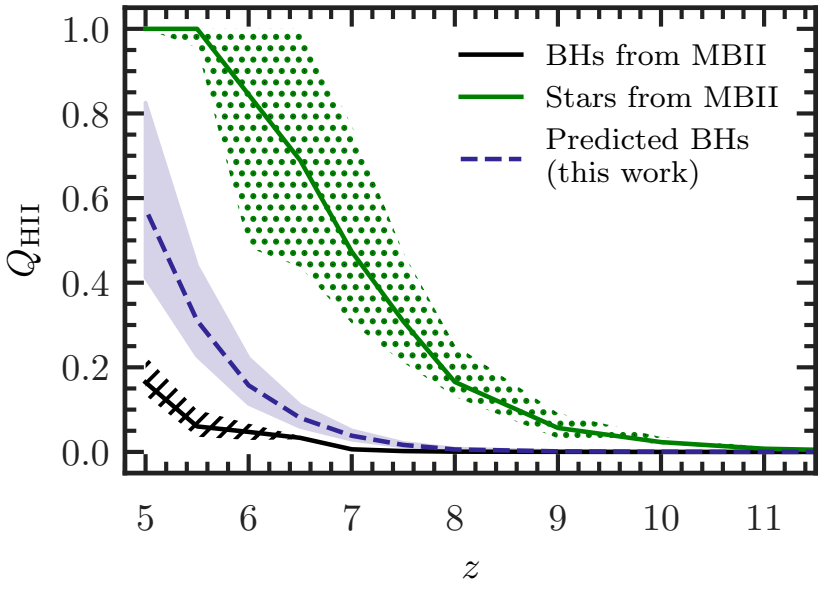

Figure 4.7: HII volume filling factor $Q_{\rm HII}$ as a function of redshift z. The lines refer to a case in which the reionization process is driven by MBII seeded BHs (solid black line in hatched area), stars in MBII (green line in dotted area) and BHs populated by our neural network (dashed blue line). The upper and lower limits refer to clumping factors $C = 1$ and $C = 10$, respectively.

sample of bright QSOs (yellow hatched area), suggesting that the contribution to the ionizing budget of our faint population is significant, although it should be noted that this is an upper limit as we have populated every galaxy with a BH. Our predicted $\bar{\varepsilon}$ is however below the inferred contributions from stars, as shown in the non-parametric model inferred by Mason et al. (2019a) from the CMB optical depth, dark Lyα and Lyβ pixels and hydrogen neutral fraction constraints from Lyα observations. Our $\bar{\varepsilon}$ overlaps with the Mason et al. (2019a) model at $z \lesssim 6$.

4.3.3 Impact on the Reionization Process

The final question we address in our study is whether such a population of faint BHs could have a significant impact on the EoR. While we plan to run simulations as those presented in Eide2018 and Eide2020 including these faint BHs, here we limit the analysis to a simpler approach. We calculate the filling factor $Q_{\rm HII}$ of ionized hydrogen (HII) as (Madau et al. 1999):

$$\frac{dQ_{\rm HII}}{dt} = \frac{f_{\rm esc}\bar{\varepsilon}}{\bar{n}_{\rm H}} - \frac{Q_{\rm HII}}{\bar{t}_{\rm rec}}, \tag{4.3}$$

where $f_{\rm esc}$ is the escape fraction of ionizing photons, $\bar{\varepsilon}$ is the volume averaged ionizing emissivity, $\bar{n}_{\rm H}$ is the average cosmic hydrogen number density and $\bar{t}_{\rm rec} = (C\bar{n}_{\rm H}\alpha(T))^{-1}$ is

the recombination time, for which we assume a clumping factor $C = 1, 5, 10$ and a case-A recombination coefficient α at $T = 10^4\,\mathrm{K}$. We calculate Q_{HII} for the MBII BHs, as well as for those seeded by our neural network, assuming $f_{\mathrm{esc}} = 1$ for both. As a comparison, we also calculate Q_{HII} for the stars of MBII, assuming $f_{\mathrm{esc}} = 0.15$ as in Eide2018 and Eide2020. These are shown in Fig. 4.7. For $C = 5$ we find that the population of mainly faint BHs seeded with our neural network results in a reionization history in which the BHs have a central, albeit not dominant, role, reaching $Q_{\mathrm{HII}} > 0.15$ (0.5) at $z = 6$ (5). This is in stark contrast to the massive BHs of MBII, which reside only in the most massive galaxies and yield $Q_{\mathrm{HII}} < 0.05$ (0.2) at the same redshifts. As expected, the stars dominate the reionization process, producing $Q_{\mathrm{HII}} \sim 1$ already at $z = 6$. Finally, we should note again that the contribution from the network generated BHs should be regarded as an upper limit, as not every galaxy is in reality expected to host an active BH. We defer to future work a refinement of this approach.

4.4 Discussion and conclusions

In the cosmological hydrodynamical simulation MassiveBlack-II (Khandai et al., 2015), galaxies with a halo mass in excess of $M_{h,\mathrm{seed}} = 5 \times 10^{10} h^{-1}\ \mathrm{M}_\odot$ are populated with seed black holes (BHs) with $M_{\mathrm{BH,seed}} = 5.5 \times 10^5 h^{-1}\ \mathrm{M}_\odot$. While this prescription assures that the BH population has physical properties consistent with observations at $z \lesssim 6$, a different seeding procedure, with BHs hosted also in smaller galaxies, might have a strong impact, among others, on the role played by BHs in the reionization process of the intergalactic medium and the related 21 cm signal. To investigate this in more detail, we have trained a neural network using the properties of galaxies harboring BHs at $z = 6$. This network allowed us to mock BHs in all galaxies down to the resolution limit of the simulations at all redshifts. By design and through training, the network replicates the properties of the pre-existing BHs in the simulation.

Our network predicts the BH masses and accretion rates of existing BHs with great precision ($> 99\%$ and $> 93\%$, respectively). Interestingly, we mock BHs with masses below the MBII seed mass when applying the network to all galaxies, also those with halo masses below $M_{h,\mathrm{seed}}$. This should not be interpreted as a mere extrapolation of the seeding procedure to lower masses, as our predictions of M_{BH} and $\dot{M}_{\mathrm{BH}}$ are constrained in 23 dimensions with a high accuracy (e.g. the predictions of the mass function in Fig. 4.4 where BHs are lighter at higher z). In fact, a galaxy with $M_h < M_{h,\mathrm{seed}}$ may still share up to 22 other parameters with galaxies hosting BHs in MBII, and thus be tightly constrained in these other dimensions. We find that removal of one parameter, including M_h, from our network did not lead to a significant deterioration of its predictions. Similarly, not a single one of the input parameters provides predictions as accurate as the full network. The exercise of removing parameters from the network, nevertheless, highlighted that the stellar mass of the galaxy, M_*, is the most important parameter. Alone, it can predict the BH mass with an accuracy of 0.88, while M_h has an accuracy of ~ 0.80. It is harder to infer the effect of the velocity dispersion. From the well-known M_{BH}–σ^* relation (Ferrarese

& Merritt, 2000) we expect the velocity dispersion to be a dominant parameter, but we cannot directly infer its role as it is not a single input to our network, but it is rather decomposed along each coordinate axis r, ϕ and θ.

While we took great care in the training of the network, its performance is still somewhat limited by the size of the training sample, both in terms of number of objects and range of masses covered. Furthermore, our network was trained on $z = 6$ galaxies hosting a BH, as only at that time does MBII produce a sizable population of BHs. This situation would improve by adopting larger and/or higher resolution simulations, such as BlueTides (Feng et al., 2015) or Illustris TNG300 (Nelson et al., 2018), or employ simulations specifically designed for this task. Nevertheless, we found that the network predicts the properties of the majority of BHs at all redshifts with high accuracy, indicating also that there is no significant evolution in the relation between the environment and the BHs' properties, in line with Huang et al. (2018). However, our slight deficiency of brighter BHs at $z > 6$ (and surplus at $z < 6$) points to these being formed more efficiently at early times (see also e.g. DeGraf et al. 2012, 2015). Our results suggest that a galaxy at $z > 6$ with properties identical to those of one at $z = 6$ is more likely to host a brighter BH.

This slight deficiency of the brightest BHs at lower z in turn ensures a perfect match at $z = 6$ to the recent LF of Kulkarni et al. (2019), and a perfect match at $-15 > M_{\rm AB} > -17$ to the LF of Giallongo et al. (2015). The most interesting feature of our results is however the large population of faint, $M_{\rm AB} > -15$, BHs. Such a population is entirely possible, as the pre-existing BHs (and the combined contributions from other energetic X-ray emitting sources) in MBII are unable to account for more than a few per cent of the unresolved X-ray background (Ma et al., 2018b), leaving ample margin for a higher contribution at high redshift.

This predicted population of BHs is unable to drive EoR alone, but it may play an important role nevertheless. Our mocked BHs do not yield enough ionizing photons to fulfil the constraints on the ionizing budget calculated from observational constraints by Mason et al. (2019a). However, our emissivites are an order of magnitude larger than those inferred from integrating the LF of the brighter QSOs of Kulkarni et al. (2019). Our BHs leave a significant imprint on the HII volume filling factor, which at $z = 5$ ranges from $Q = 0.41$ with a clumping factor $C = 10$ to $Q = 0.83$ with $C = 1$. The existing BHs in MBII can at best yield $Q = 0.23$ with $C = 1$, but while this population satisfy the bright end of the LF down to $z > 2$, it does not include the fainter population that our network predicts. Our population of mocked BHs is neither negligible, nor is it as dominating as the one of Madau & Haardt (2015). Further work is needed to investigate whether they will induce an extended HeII reionization epoch as observations imply (Worseck et al., 2016, 2019) without providing undue heating (see e.g. D'Aloisio et al., 2017; Garaldi et al., 2019). We plan to investigate this more in detail with numerical simulations following the work of Eide2018 and Eide2020.

A more prominent population of high-z, small mass BHs could also have an important impact on the 21 cm signal from neutral hydrogen in the IGM, by partially ionizing and heating the gas prior to full reionization (e.g. Madau et al. 1997).

Our conclusions can be summarized as follows.

- We train a neural network on properties of BH hosting galaxies at $z = 6$. For our training sample, this predicts the mass, $M_{\rm BH}$, and accretion rate, $\dot{M}_{\rm BH}$ of BHs with an accuracy $> 99\%$ and $> 93\%$, respectively. These properties at other redshifts are also predicted with high precision.

- $M_{\rm BH}$ and $\dot{M}_{\rm BH}$ are predicted with the most relevant single parameter, the stellar mass M_*, with an accuracy of 88% and 75%, respectively. Removing M_* degrades the network to accuracies of 98.6% and 88.8%. The predictions of our network are robust, even when single parameters are ill-defined.

- The neural network is slightly less effective at predicting the brightest and most massive BHs at $z > 6$, and conversely predicts a population of slightly brighter BHs at $z < 6$. This points to a decrease in BH formation efficiency with decreasing z.

- Populating all galaxies with a nuclear BH, we predict a substantial population with mass below that of the seeds at all redshifts. This results in a LF at $z = 6$ with a knee at $M_{\rm AB} = -15$ and a lack of turnover at least down to $M_{\rm AB} = -5$.

- Our predicted population of BHs can contribute significantly to H reionization, yielding a Universe in which H is $\sim 15\%$ ionized by BHs at $z = 6$ for a clumping factor of 5. The bright BHs alone, which are well reproduced by MBII, predict instead a Universe that is only $\sim 5\%$ ionized at the same redshift.

Chapter 5

The Conclusion

We started out by asking two simple, but fundamental questions,

1. What was the temperature of the early Universe?
2. How did the Universe become ionized?

We are now in a position where we can summarize our findings.

The Cosmic Dark Ages began a few hundred thousand years after the beginning of times, and concluded with the Epoch of Reionization roughly a billion years later. During this epoch, the primordial gas that remained in between the galaxies became heated and ionized. The problem we sought to investigate in this thesis was how early sources of radiation could have caused this transformation, and what their signatures would be. These questions are particularly relevant as current and next generation radio telescopes begin to peek into the Dark Ages and need predictions that their observations can be compared to.

In this work, we did not consider the very first sources of light and heat such as the first stars (PopIII objects) or black holes made from primordial gas that could have directly collapsed (direct collapse BHs) during the Dark Ages. Although they were powerful radiators, there were likely few of them. In our work, we have therefore employed sources of ionizing radiation known from low redshifts, but in a high redshift environment. We chose four ionizing source types—stars, nuclear BHs, XRBs and diffuse radiation from the supernova-heated ISM. We identified these sources in the hydrodynamic structure formation simulation MassiveBlack-II (MBII, Khandai et al., 2015) whose parameters were tuned to provide a simulated present-day Universe. We then simulated reionization by post-processing said simulation with the three-dimensional, multifrequency Monte Carlo-based radiative transfer code `CRASH` (Ciardi et al., 2001; Maselli et al., 2003, 2009; Graziani et al., 2013; Hariharan et al., 2017; Graziani et al., 2018; Glatzle et al., 2019).

The source types and their properties were determined by us through a process where we used identified halos and either assigned spectra and luminosities to stellar and black hole particles or assigned galaxy-wide radiative properties based on the halos' physical environment. We chose an approach where we used constraints and models from present-day theory and observations so as to both test their validity at high redshift but also to

study the reionization history these models would imply. The only free parameter for us to tweak was the debated and unknown *escape fraction of ionizing radiation*, $f_{\rm esc}$. We chose two models for it, one in which $f_{\rm esc}$ evolves with redshift and follows a standard power-law model with parameters inferred from joint observations of reionization signatures and tuned to scale the stellar emissivity to reach reionization by $z = 6$, and another in which we have a constant escape fraction of 15%. One of our findings was that an evolving escape fraction provides the required attenuation at $z < 6$ to evolve the fraction of neutral hydrogen in a way that matches observational constraints. An important caveat here is that we are unable to resolve small, dense systems which are efficient absorbers, the Lyman Limit Systems (LLS), which themselves provide attenuation (see e.g. Madau et al., 1999).

The stellar properties were determined by the physical characteristics of the stellar particles in the hydrodynamical cosmological structure formation simulation, with lower age, lower metallicity and higher mass leading to higher ionizing emissivity. We obtained the stellar spectra and luminosities by looking them up in the population synthesis model `BPASS` (Eldridge & Stanway, 2012) using the physical properties of the stellar particles in MBII. We found that the numbers and luminosities of the stars by far dominated the total cosmological ionizing budget over BHs, XRBs and the ISM. The stars can therefore be the sole drivers of hydrogen reionization. Their spectra is of a nature which leaves little ionizing radiation to escape beyond the ionized regions they carve out in the IGM. A prime signature of a star-only driven hydrogen reionization would be the absence of heating of the neutral IGM.

Although stars are the undisputed driver of hydrogen reionization, BHs have long been theorized to play a central role. We therefore examined whether we were artificially down-playing their importance, as in our radiative transfer setup, we only consider a reionization scenario where none but the most massive galaxies hold a BH. This choice was dictated by the seeding procedure adopted in the hydrodynamic simulation we employed.

To investigate this point, we chose to generate a neural network to predict BH masses and accretion rates given the full set of galactic properties we had. An alternative to this approach would have been to cherry-pick a few physical galactic parameters to create a BH-galaxy relation. The strength of using a neural network is that we leave it to the network training process to determine the importance and non-linear relationships between the parameters. Moreover, we benefited greatly from the agility of the network. We could reliably make sane predictions of BH properties when we populated galaxies whose masses were too low for them to have been given a BH in the hydrodynamic simulation. This was made possible by the high dimensionality and the non-linearity of the network. When the mass-parameter was outside of the range it had been trained on, the network had to extrapolate its predictions in this one dimension. However, it still had 22 other properties that were well defined when making its predictions.

We predict a population of faint and light BHs to reside in the nuclei of lighter galaxies. This population is missing from hydrodynamic simulations, and our predictions on their impact on reionization may be considered an upper bound of their importance. We did not find that they could dominate the ionizing budget, but that their numbers were sufficiently large to alone ionize more than 40% of the Universe by $z = 5$. We also made the interesting

discovery that BHs grow more efficiently at higher z.

Returning to our reheating and reionization simulations, we did nevertheless find that the smaller population of bright BHs in the most massive galaxies leave unambiguous signatures on the IGM. We expect our predicted fainter population to boost these signatures further. The most important finding is that only BHs can drive helium reionization, and their telltale signature is HeIII regions. Furthermore, BHs have spectra which make them both efficient ionizers but also efficient heaters of gas outside HII regions.

The role of the XRBs and the ISM serve to illustrate the impact of two source types which have radically subdominant ionizing outputs compared to the stars, but which nevertheless hallmark their presence by leaving two distinctly different signatures. The XRBs have spectra that peak at keV scales and are poor heaters, but nevertheless provide rather uniform, long range partial ionization of the IGM outside HII regions. The population of XRBs arises too late and is too few in numbers for it to provide the wealth of radiation that eventually could redshift and impact the IGM at a later time, as questioned by Fialkov et al. (2014b). The IGM is already well underway towards being fully ionized and hot when the hard XRB radiation emitted during the early Cosmic Dawn finally is redshifted to wavelengths where it is likely to interact with neutral hydrogen and helium. The ISM, on the other hand, has a spectrum much softer than the XRBs, but one that still is fainter at stellar UV frequencies. Its spectrum can be described as one that is providing swaths of hard UV photons and soft X-rays, efficiently penetrating deep into neutral HI gas, heating and partially ionizing it.

Our XRB and ISM models can nevertheless be considered useful to predict the impact of an unspecified source type with a spectral signature that is similar to either of our spectrally distinct source types. Our exercises are therefore not futile should there be a dissonance between our theoretically modelled sources and the actual ones at high redshifts. Nevertheless, this is a caveat of our approach, we have assumed the high redshift Universe to have sources with spectral signatures similar to those we are familiar with from observations at low redshifts.

The precision and reliability in the ionizing RT approach we have chosen mean that our set of simulations are treasure troves for future studies that are sensitive to the accuracy of the RT approach. We have already employed the suite of simulations to uncover that our X-ray sources during Cosmic Dawn contribute less than a few percent to the diffuse present-day X-ray background in Ma et al. (2018b), the prospects of cross correlating [OIII] radiation at high redshifts with 21 cm radiation in Moriwaki et al. (2019), in examining the potential use of the hyperfine spin-flip 3.5 cm transition of $^3He^+$ to study He at high redshifts in Khullar et al. (2020), in studying the morphological properties of the phase-changing IGM in Busch et al. (2020) and in an upcoming study (Ma et al., in prep.) where we examine the detailed HI 21 cm signatures our sources imprint as they heat and ionize gas during the Cosmic Dawn.

The story of how the Cosmic Dark ages progressed trough the epochs of Cosmic Heating and Reionization can be summarized in a most precise, scientific way: it depends. It depends on the source types existing and playing a significant role during the process. It depends on their detailed spectra. It depends on the amount of galactic ionizing radiation

that is allowed to escape into intergalactic space. It depends on the choice of model for star formation and structure assembly. It depends on whether faint BHs are present in all galaxies during Cosmic Dawn. It depends on whether present observations, constraints and models can be trusted to the point where you let them guide you in your search for the history of the Universe's first billion years.

Should stars be present, but not in binary pairs, BHs only exist in the most massive galaxies, XRBs exist, and SNe shock heat the ISM for it to radiate diffusively, we may have the following answers to the two questions asked in chapter 1 and which were repeated at the beginning of this chapter. The temperature of the IGM is either one of following three. It can be untouched by ionizing radiation, cooling adiabatically, coldest to the point of being a few degrees above absolute zero in the most underdense regions, and be slightly warmer where structures are forming. Or it can be fully ionized by the radiation from the early second generation stars, and be in the process of cooling off and being ionized anew from the initial $\sim 20,000$ K it was heated to when the first front of stellar ionizing radiation blew past it. BHs will fully ionize helium and heat the gas in their vicinity further $20,000$ degrees. Or, finally, the gas may be impinged upon by the hard UV- and X-rays of the faint ISM or XRBs, and be heated tens and hundreds of degrees, efficiently casting it to a form where it would be seen as glowing hot, not CMB-absorbingly cold in observations of the redshifted HI 21 cm line from the Cosmic Dawn. As for the second question—the stars ionized the intergalactic hydrogen and BHs stood for helium reionization.

We have presented a cosmological reheating and reionization history with clear and distinct source signatures that may guide observational efforts. The job falls to the present-day and future Earth-bound radio telescope arrays capable of recording HI 21 cm radiation emitted at high redshifts, such as LOFAR (van Haarlem et al., 2013), MWA (Tingay et al., 2013), PAPER (Parsons et al., 2010), GMRT (Swarup, 1991), HERA (DeBoer et al., 2017) and SKA, or to single all-sky averaging dipole experiments such as SARAS (Singh et al., 2017), BIGHORN (Sokolowski et al., 2015), SCI-HI (Voytek et al., 2014), LEDA (Greenhill & Bernardi, 2012) and EDGES (Bowman & Rogers, 2010).

However, the days for Earth-based radio and optical astronomy are numbered. Future surveys will have to cope with the arrival of mega constellations of low Earth orbit satellites such as Starlink from SpaceX, OneWeb or KLEO who operate either optically or at wavelengths shorter than 2.5 cm, or Lynk which will operate at cell phone frequencies interfering with high-redshift 21 cm surveys. The obvious solution is to go to space, which is what the People's Republic of China now do with their Chang'e 4 lunar mission hosting NCLE - The Netherlands-China Low Frequency Explorer, pursuing a path that seeks to establish a constellation of radio interferometers orbiting the Moon. There is movement at both NASA, which has previously shown interest in DARE - the Dark Ages Radio Explorer (Burns et al., 2012), and now in DAPPER - the Dark Ages Polarimeter Pathfinder, and at ESA, which previously has twice rejected a similar mission, DEX - the Dark Ages eXplorer. The latter serves as an illustration of a situation where there is political will but unfortunately also hesitation in the scientific community.

However, the sheer magnitude of the undertakings needed to study the Dark Ages from space calls for international collaboration. The infrastructure needed to do lunar radio

astronomy requires tremendous technological and financial advances to be made, but such challenges are not unknown to the radio astronomy community. Going off Earth means that the fruits of such advances would not be only for the scientific community to harvest, but for humanity as a whole to thrive and grow on. The necessity for collaboration in such endeavours may inadvertently foster a lasting peaceful atmosphere and instigate a deep sense of shared, common responsibility among the parties. History has taught us that similar efforts can withstand even the most volatile geopolitical storms, with the International Space Station being a prime example. The far side of the Moon would be the ideal place to study the Cosmic Dark Ages, and could prove to be the catalyst that drives humanity peacefully into space.

And so this thesis ends—with an outlook to the Moon, Mars and the stars.

Appendix A

The Epoch of Cosmic Heating

A.1 Convergence

To check the convergence of our results, we run two simulations including all source types which differ only in the number of photon packets emitted per source, i.e. $N_\gamma = 10^4$ and $N_\gamma = 10^5$.

In Fig. A.1 we show the fraction of cells at $z = 12$ with a given T, $x_{\rm HII}$, $x_{\rm HeII}$ and $x_{\rm HeIII}$, together with the relative difference between the results of the two simulations, i.e. $D/D_{\rm ref} - 1$, where $D = T, x_{\rm HII}, x_{\rm HeII}, x_{\rm HeIII}$ and $D_{\rm ref}$ represents the values obtained with the highest value of N_γ. More than 50% of the cells in the $N_\gamma = 10^4$ simulation are very well converged, with differences at the percent-level, except for $x_{\rm HII} = x_{\rm HeII} \sim 10^{-3}$ where the results deviate up to $\sim 20\%$. For the temperatures, the differences are less than 5% for the 67% of the cells that have 10 K $< T \leq 10^4$ K. Large differences are observed only in a handful of cells at high temperature or very low ionization fraction $x_{\rm HII} = x_{\rm HeII} < 10^{-5}$, which we set as the lower limit in this work.

A.2 Lightcones

We construct lightcones inspired by the methods applied by Mellema et al. (2006b), Datta et al. (2012) and Giri et al. (2018). Our approach though needs to take into account the non periodic boundary conditions of our simulations. We create a path $\Delta l(t_i) \equiv \Delta l_i$ corresponding to the light travel distance $c\Delta t_i = k_i \Delta x_i$ covered by a ray travelling through k_i cells of the volume, each with sides of physical length

$$\Delta x_i = \frac{1}{1+z_i} \frac{100 h^{-1}}{256} \text{pMpc}, \tag{A.1}$$

from one `CRASH` output i to the next, $i+1$, having a difference in cosmological age Δt_i. We hence assume the time steps to be sufficiently small to neglect the evolution between them when calculating the path Δl_i. We linearly interpolate between each output, i.e. the k_i cells that are covered between two `CRASH` snapshots have contributions from both snapshots.

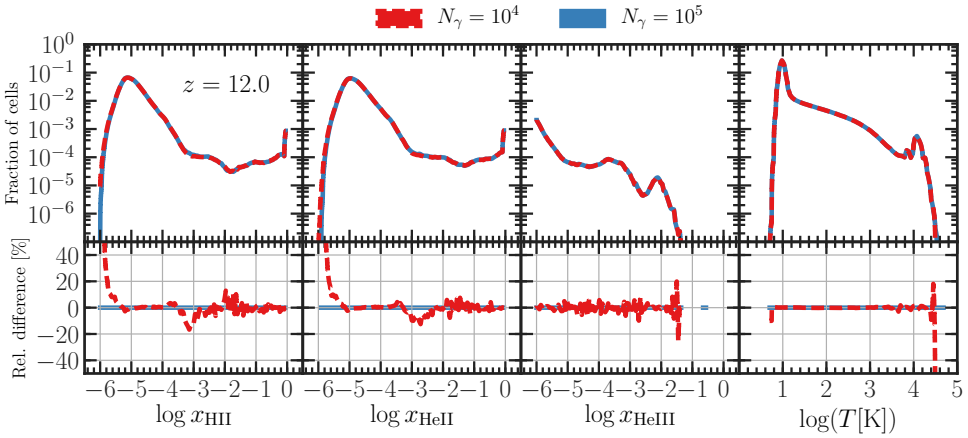

Figure A.1: *Upper panels*: from left to right, fraction of cells at $z = 12$ with a given $x_{\rm HII}$, $x_{\rm HeII}$, $x_{\rm HeIII}$ and T for simulations with $N_\gamma = 10^4$ (red dashed lines) and 10^5 (blue solid lines). *Lower panels*: relative difference in percent with respect to the case with $N_\gamma = 10^5$.

The position along the k_i cells that make up the path between two snapshots can be written as l_p, where $p = 0, \ldots, k_i - 1$. For each point l_p, which corresponds to a position in both physical (x_p, y_p) and redshift (z_p) space, we also have a third spatial dimension, $\mathbf{c}_p(l_p)$, which is plotted as the y-axis in the lightcones. It has contributions from the two snapshots at times t_i and t_{i+1}, weighed by the position $(k_i - 1 - p)/(k_i - 1)$ and $p/(k_i - 1)$, respectively, along the path between them. For the ionization fraction we have

$$\mathbf{c}_p(l_p) = \frac{1}{k_i - 1}\left[(k_i - 1 - p)\,\mathbf{x}_{\rm HII}^{t_i}(l_p) + p\mathbf{x}_{\rm HII}^{t_{i+1}}(l_p)\right], \tag{A.2}$$

where we write the ionization fractions ${\mathbf{x}_{\rm HII}}^{t_i}(l_p)$ from the two contributing snapshots as vectors to indicate that we only use the cells along the third spatial dimension at this point in (z_p, x_p, y_p)-space.

There are various possible approaches for deciding the paths Δl_i. Instead of choosing random paths, which could overlap, we instead try fixed patterns, and attempt at covering as much of the simulations volumes as possible in the lightcones. For this purpose, we choose a continuous whisk broom scanning approach (see e.g. Schaepman, 2009) with decreasing horizontal scan width and a several pixels wide (slightly decreasing) vertical scan offset between each horizontal scan. Both are reset to maximum widths once we reach the vertical boundary. This way we prevent, as far as possible, the same objects at the same cosmological ages to appear in the lightcones. We only use the interior 236 cells (out of a maximum of 256) along each scanning axis to prevent boundary effects on the lightcones. The parameters of this approach (the number of pixels to decrease the scan width and offset, vertical scan offset) must however be tuned to obtain a satisfying result. The approach is thus optimal for a visualization, rather than for an objective quantification, of the temporal progression.

Appendix B

The Epoch of Reionization

B.1 Loss of high energy photons

While energetic photons are important for partial ionization and heating, their mean free path is very long and thus, with our assumption of non-periodicity in the RT, they could easily escape from the simulation box and be lost.We do however account for and track the redshifting of each photon as it propagates through the box. In the following we show that this is not expected to have a significant impact on the results presented in this chapter.

In Fig. B.1 we plot at which redshift, $z_{\rm req}$, a photon of energy E needs to be emitted in order to be redshifted to 100 eV (and thus be easily absorbed) by a given redshift z. As a reference, to reach 100 eV by the end of the simulation at $z \sim 5$, a photon of 300 eV should have been emitted at $z \sim 17$. As the IGM has already been partially or fully ionized well before $z = 5$ by less energetic photons, this suggests that photons above a couple of hundred eV need times longer than those available in the simulation to possibly play a relevant role.

In Fig. B.2 we plot the distribution of the mean free path (MFP) of cells in the simulation with all source types and $f_{\rm esc}(z)$ at $z = 7$. When only highly neutral cells (i.e. with $x_{\rm HII} \leq 0.9$) are considered (top panel), the distributions are strongly peaked and shift to larger mfp with increasing photon energy. For all photons with $h_{\rm P}\nu < 100$ eV the MFP is smaller than the box dimension. When only highly ionized cells (i.e. with $x_{\rm HII} > 0.9$) are considered (bottom panel), instead, the distributions are more complex, as the presence of He (and its ionization state) becomes more relevant. We then see that the MFP of photons below the ionization threshold of HeII becomes even larger than the box size, as HI and HeI are almost fully ionized, and the corresponding distributions are much wider, reflecting the strong dependence on the detailed ionization state of the various components of the gas. Conversely, higher energy photons have distributions which still resemble those in the left panel, with a strong peak, albeit shifted towards larger values of the MFP due to the reduced absorption from HI and HeI.

In Fig. B.3 we show the redshift evolution of the MFPs for photons of different energies. These have been calculated by taking the median of the MFPs in cells where hydrogen is

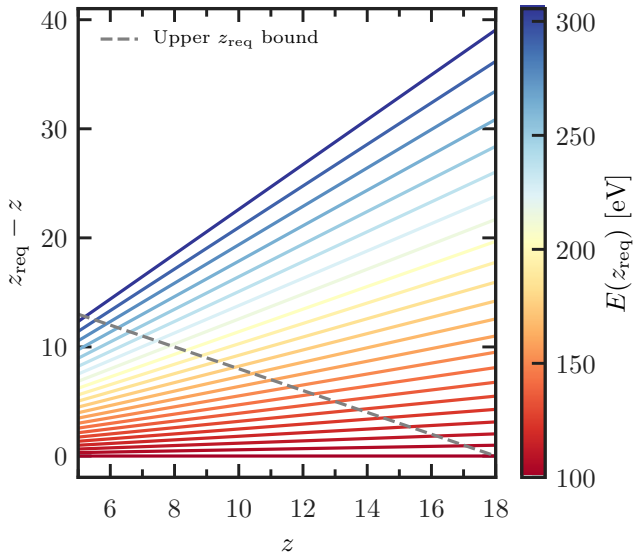

Figure B.1: Redshift at which a photon with energy indicated by the color bar needs to be emitted in order to be shifted to 100 eV by redshift z indicated in the x-axis. Photons occupying the area above the grey dashed line would need to be emitted at redshifts above those covered by our simulations.

highly neutral ($x_{\rm HII} \leq 0.9$, first panel) as well as where it is highly ionized ($x_{\rm HII} > 0.9$, second panel); by taking the mean of the MFPs in all cells (third panel) and finally by inferring the MFP from the volume averaged ionization fractions $\langle x_{\rm HII} \rangle$, $\langle x_{\rm HeII} \rangle$ and $\langle x_{\rm HeIII} \rangle$. The results in the highly neutral IGM reflect what was observed in the previous figure, i.e. the MFP smoothly increases with increasing photon energy. All photons with energies below a couple of hundred eV have MFPs shorter than the box size at all redshifts. Conversely, for the highest energy photons this happens only at high redshift. For example, the MFP of a 1 keV photon becomes larger than $100h^{-1}$ cMpc at $z < 8$. This increase in MFP with decreasing redshift observed for all photon energies is associated to the declining gas number density with the cosmological expansion and is proportional to $1/z^3$.

In highly ionized regions (second panel), the MFP of photons with energies below 54.4 eV becomes much longer due almost lack of HI and HeI, and once reionization is well under way, it becomes even larger than the box size. For these photons there is also a stronger dependence on redshift, with a more rapid increase observed from $z \sim 9$, when $x_{\rm HII} > 10^{-1}$. Photons with energies above 54.4 eV, instead, have a behaviour similar to the one observed in the previous plot, but the MFPs can be more than one order of magnitude larger. This difference means that even the highest energy photons are sensitive to the presence (or absence) of HI (and HeI) even though the ionizing cross section of hydrogen $\sigma_{\rm HI}$ at such energies is extremely small (as $\sigma_{\rm HI} \propto (E/13.6\,{\rm eV})^{-3}$).

The volume averaged MFP (irrespective of HI ionization state, as plotted in the third panel) displays how the MFPs transition from being dominated by those we found in the

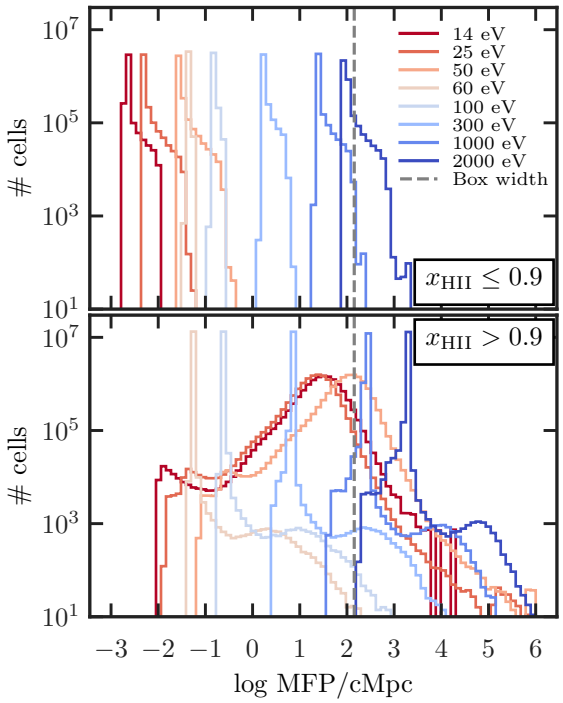

Figure B.2: Number of cells with a given mean free path in the simulation with all source types and $f_{\rm esc}(z)$ at $z = 7$. The distributions refer to photons of different energies (as indicated by the colors), while the vertical grey dashed line indicates the box dimension. From top and bottom the panels refer to cells with $x_{\rm HII} \leq 0.9$ and $x_{\rm HII} > 0.9$, respectively.

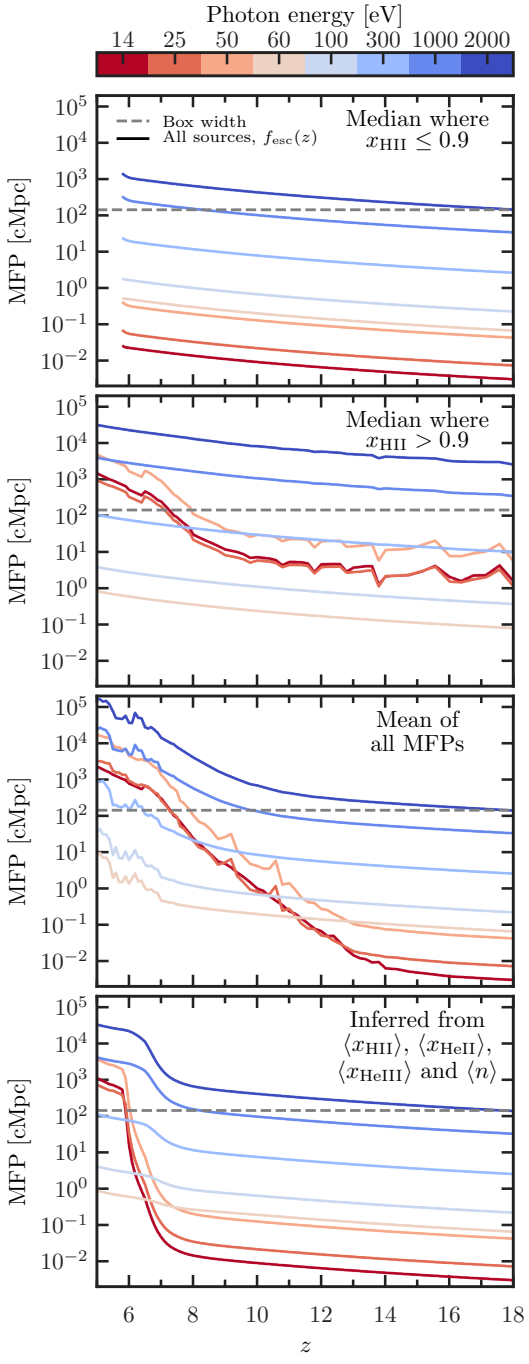

Figure B.3: Redshift evolution of the mean free path (derived from the statistics of the MFPs in each simulation cell) of photons of different energies (as indicated by the line colour) in the simulation with all source types and $f_{\rm esc}(z)$. The grey dashed lines refer to the length of the simulation box. From to to bottom the panels refer to: the median of the MFPs in cells with $x_{\rm HII} \leq 0.9$; the median in cells with $x_{\rm HII} > 0.9$; the volume average of all the MFPs; and the MFPs one would obtain by calculating it directly from the volume averaged number density $\langle n \rangle$ and ionization fractions $\langle x_{\rm HII} \rangle$, $\langle x_{\rm HeII} \rangle$ and $\langle x_{\rm HeIII} \rangle$of the simulation.

neutral HI at high z to the much longer ones in HII regions at lower z. This increase is most dramatic and extended for HI and HeI ionizing photons, whose MFPs increase by more than five orders of magnitude between $z \sim 14$ and the end of the simulation. For higher energy HeII ionizing photons we see the same uptick in MFPs at $z < 6$, when HeII ionization becomes significant.

Finally, in the last panel we show the MFPs one would obtain by calculating them directly from the volume averaged ionization fractions $\langle x_{\rm HII} \rangle$, $\langle x_{\rm HeII} \rangle$ and $\langle x_{\rm HeIII} \rangle$ and number density. The MFPs display the same behaviour as in the third panel, albeit with a much shorter transition period at $z < 7$ from being dominated by MFPs in the neutral HI to the ionized HII at $z = 6$.

We conclude by observing that even very high energy photons have MFPs similar to our box length as long as some HI is still present. It should be highlighted, though, that photons above a couple of hundred eV need times longer than those available in the simulation to possibly play a relevant role.

B.2 Reionization timing

We define the reionization redshift $z_{\rm reion}$ of the cell as the redshift at which the cell has reached a hydrogen ionization fraction $x_{\rm HII}$ larger than a threshold value, $x_{\rm HII}^{th}$. In our reference case we adopt $x_{\rm HII}^{th} = 0.9$. In Fig. B.4 we show how $z_{\rm reion}$ changes when we lower this threshold 0.1 for the simulation with all source types and $f_{\rm esc} = 15\%$. We find that $z_{\rm reion}$ increases by as much as $\Delta z_{\rm reion} \sim 2$ in the vicinity of the sources, while it remains unaltered for the majority of the IGM, with $\Delta z_{\rm reion} < 0.1$ for 50% of the gas, and $\Delta z_{\rm reion} < 0.4$ for 90% of the gas.

In Fig. B.5 we show how $z_{\rm reion}$ changes with different source types for a reference value of $x_{\rm HII}^{th} = 0.9$. We find that reionization in the vicinity of BHs happens earlier, by more than a factor of 0.2 in redshift, whereas XRBs and ISM have a very small impact. Although some of the features observed in all these panels closely resemble Monte Carlo noise, we have verified that this is not the case for all cells and they are instead related to the complexity of the multi-frequency radiative transfer. We highlight one such region with a rectangle in the 'Stars,XRBs' panel of the figure. Indeed some gas pockets experience earlier reionization due to the longer mean free path of the high energy photons emitted by these sources, but at the same time, because such photons ionize less efficiently, for some cells we observe a delayed reionization. The behaviour of the simulation including all source types reflexes that observed in the simulation with stars and BHs, with reionization occurring slightly earlier ($\Delta z_{\rm reion} \approx 0.1$) in the vicinity of BHs.

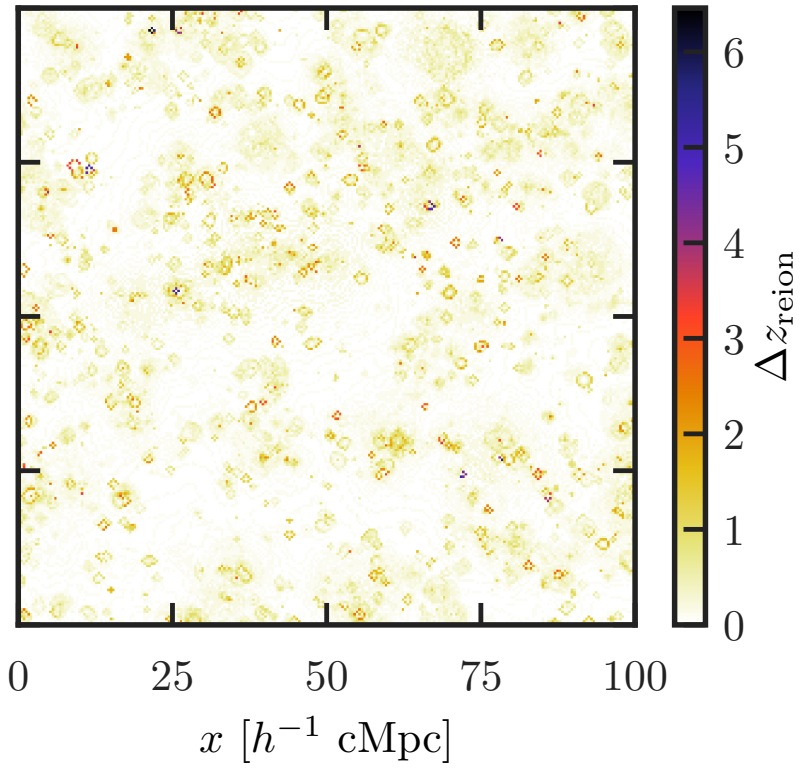

Figure B.4: Map showing the difference in the timing of reionization, $z_{\rm reion}$, for a simulation with all source types and $f_{\rm esc} = 15\%$ when we adopt an ionization threshold $x_{\rm HII}^{th} = 0.1$ with respect to the case when the threshold is 0.9.

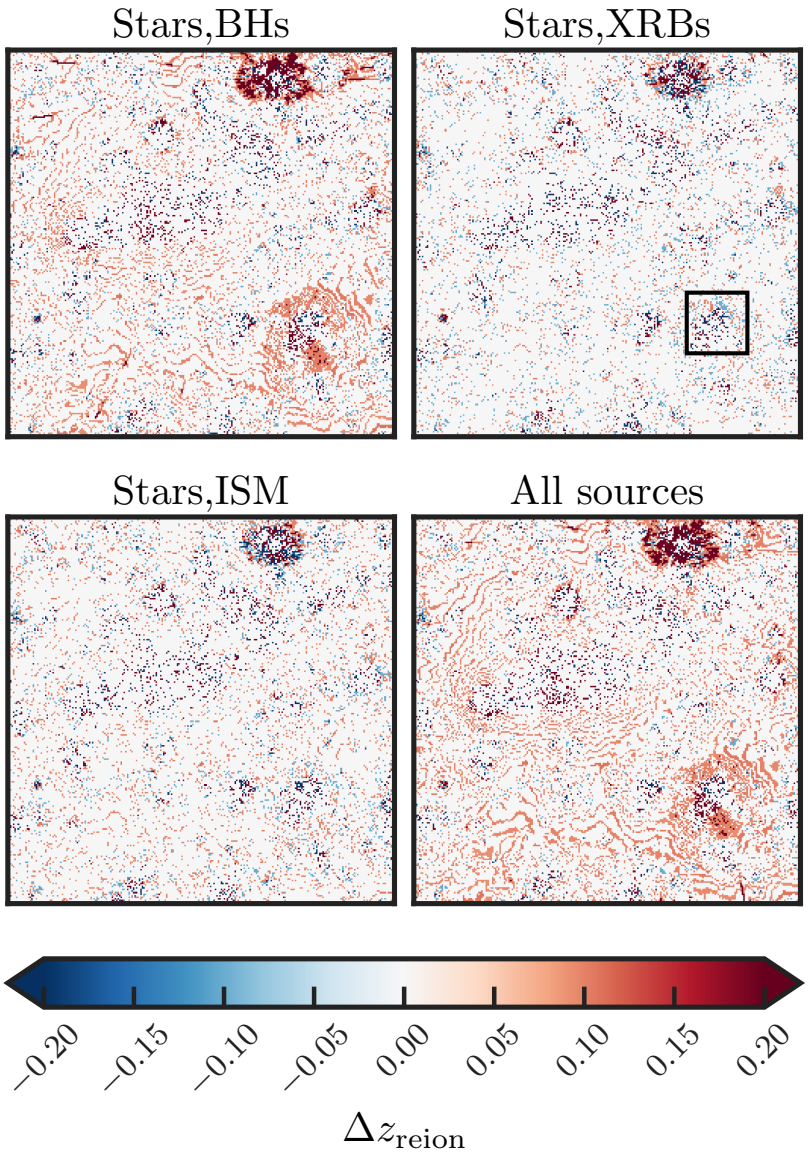

Figure B.5: Maps showing the differences in the timing of reionization, $z_{\rm reion}$, for simulations with different source types (as indicated by the labels), $f_{\rm esc} = 15\%$ and an ionization threshold $x_{\rm HII}^{th} = 0.9$, with respect to the simulation with only stars. The maps are $100h^{-1}$ cMpc wide. The black rectangle highlights a region in which the features resembling Monte Carlo noise are actually due to physical reasons (see text for more details).

Bibliography

Abel T., Haehnelt M. G., 1999, ApJ, 520, L13

Ali Z. S., et al., 2015, ApJ, 809, 61

Allen R. J., 1969, A&A, 3, 382

Arons J., McCray R., 1970, Astrophys. Lett., 5, 123

Aubert D., et al., 2018, ApJ, 856, L22

Bañados E., et al., 2018, Nature, 553, 473

Baek S., Semelin B., Di Matteo P., Revaz Y., Combes F., 2010, A&A, 523, A4

Bag S., Mondal R., Sarkar P., Bharadwaj S., Choudhury T. R., Sahni V., 2019, MNRAS, 485, 2235

Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101

Bolton J. S., Haehnelt M. G., 2007, MNRAS, 382, 325

Bolton J. S., Becker G. D., Raskutti S., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2012, MNRAS, 419, 2880

Bouwens R. J., et al., 2015a, ApJ, 803, 34

Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015b, ApJ, 811, 140

Bowman J. D., Rogers A. E. E., 2010, Nature, 468, 796

Bowman J. D., Rogers A. E. E., Hewitt J. N., 2008, ApJ, 676, 1

Burns J. O., et al., 2012, Advances in Space Research, 49, 433

Busch P., Eide M. B., Ciardi B., Kakiichi K., 2020, MNRAS,

Cen R., 1992, ApJS, 78, 341

Cen R., 2002, ApJS, 141, 211

Chang P., Broderick A. E., Pfrommer C., 2012, ApJ, 752, 23

Chardin J., Haehnelt M. G., Aubert D., Puchwein E., 2015, MNRAS, 453, 2943

Chardin J., Puchwein E., Haehnelt M. G., 2017, MNRAS, 465, 3429

Chevalier R. A., Clegg A. W., 1985, Nature, 317, 44

Chisholm J., et al., 2018, A&A, 616, A30

Ciardi B., Ferrara A., 2005, Space Sci. Rev., 116, 625

Ciardi B., Madau P., 2003, ApJ, 596, 1

Ciardi B., Ferrara A., Marri S., Raimondo G., 2001, MNRAS, 324, 381

Ciardi B., Bolton J. S., Maselli A., Graziani L., 2012, MNRAS, 423, 558

Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53

Crain R. A., et al., 2015, MNRAS, 450, 1937

Croft R. A. C., Di Matteo T., Springel V., Hernquist L., 2009, MNRAS, 400, 43

D'Aloisio A., Upton Sanderbeck P. R., McQuinn M., Trac H., Shapiro P. R., 2017, MNRAS, 468, 4691

D'Aloisio A., McQuinn M., Maupin O., Davies F. B., Trac H., Fuller S., Upton Sanderbeck P. R., 2019, ApJ, 874, 154

Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12

Dalgarno A., Yan M., Liu W., 1999, ApJS, 125, 237

Das A., Mesinger A., Pallottini A., Ferrara A., Wise J. H., 2017, MNRAS, 469, 1166

Datta K. K., Mellema G., Mao Y., Iliev I. T., Shapiro P. R., Ahn K., 2012, MNRAS, 424, 1877

Davies R. D., Fennison R. C., 1964, MNRAS, 128, 123

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dawoodbhoy T., et al., 2018, MNRAS, 480, 1740

DeBoer D. R., et al., 2017, PASP, 129, 045001

DeGraf C., Di Matteo T., Khandai N., Croft R., 2012, ApJ, 755, L8

DeGraf C., Di Matteo T., Treu T., Feng Y., Woo J. H., Park D., 2015, MNRAS, 454, 913

Degraf C., Di Matteo T., Springel V., 2010, MNRAS, 402, 1927

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33

Di Matteo T., Khandai N., DeGraf C., Feng Y., Croft R. A. C., Lopez J., Springel V., 2012, ApJ, 745, L29

Di Matteo T., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2017, MNRAS, 467, 4243

Dijkstra M., 2014, Publ. Astron. Soc. Australia, 31, e040

Dillon J. S., et al., 2014, Phys. Rev. D, 89, 023002

Eide M. B., Graziani L., Ciardi B., Feng Y., Kakiichi K., Di Matteo T., 2018, MNRAS, 476, 1174

Eide M. B., Ciardi B., Graziani L., Busch P., Feng Y., Matteo T. D., 2020a, MNRAS,

Eide M. B., Ciardi B., Feng Y., Di Matteo T., 2020b, MNRAS (acc.)

Einasto J., Joeveer M., Saar E., 1980, MNRAS, 193, 353

Einstein A., 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, pp 142–152

Eldridge J. J., Stanway E. R., 2012, MNRAS, 419, 479

Ewall-Wice A., et al., 2016, MNRAS, 460, 4320

Fabbiano G., 1989, ARA&A, 27, 87

Fabbiano G., 2006, ARA&A, 44, 323

Fan X., et al., 2000, AJ, 120, 1167

Fan X., et al., 2006, AJ, 132, 117

Fan X., et al., 2019, BAAS, 51, 121

Feng Y., Di Matteo T., Croft R., Tenneti A., Bird S., Battaglia N., Wilkins S., 2015, ApJ, 808, L17

Fernandez E. R., Zaroubi S., Iliev I. T., Mellema G., Jelić V., 2014, MNRAS, 440, 298

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Fialkov A., Barkana R., Pinhas A., Visbal E., 2014a, MNRAS, 437, L36

Fialkov A., Barkana R., Visbal E., 2014b, Nature, 506, 197

Field G. B., 1959, ApJ, 129, 525

Field G. B., 1962, ApJ, 135, 684

Field G. B., 1972, ARA&A, 10, 227

Finkelstein S. L., et al., 2019, ApJ, 879, 36

Finlator K., Keating L., Oppenheimer B. D., Davé R., Zackrisson E., 2018, MNRAS, 480, 2628

Fletcher T. J., Tang M., Robertson B. E., Nakajima K., Ellis R. S., Stark D. P., Inoue A., 2019, ApJ, 878, 87

Fragos T., et al., 2013a, ApJ, 764, 41

Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2013b, ApJ, 776, L31

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Fukugita M., Kuramashi Y., Mino H., Okawa M., Ukawa A., 1994, Phys. Rev. Lett., 73, 2176

Furlanetto S. R., Oh S. P., 2016, MNRAS, 457, 1813

Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1

Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181

Gallerani S., Ferrara A., Fan X., Choudhury T. R., 2008a, MNRAS, 386, 359

Gallerani S., Salvaterra R., Ferrara A., Choudhury T. R., 2008b, MNRAS, 388, L84

Garaldi E., Compostella M., Porciani C., 2019, MNRAS, 483, 5301

Ghara R., et al., 2020, MNRAS, 493, 4728

Giallongo E., et al., 2015, A&A, 578, A83

Ginzburg V. L., Ozernoi L. M., 1965, Azh, 42, 943

Giri S. K., Mellema G., Dixon K. L., Iliev I. T., 2018, MNRAS, 473, 2949

Glatzle M., Ciardi B., Graziani L., 2019, MNRAS, 482, 321

Gnedin N. Y., 2014, ApJ, 793, 29

Gnedin N. Y., Abel T., 2001, New Astron., 6, 437

Graziani L., Maselli A., Ciardi B., 2013, MNRAS, 431, 722

Graziani L., Salvadori S., Schneider R., Kawata D., de Bennassuti M., Maselli A., 2015, MNRAS, 449, 3137

Graziani L., Ciardi B., Glatzle M., 2018, MNRAS, 479, 4320

Greenhill L. J., Bernardi G., 2012, arXiv e-prints, p. arXiv:1201.1700

Greig B., Mesinger A., Pober J. C., 2016, MNRAS, 455, 4295

Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239

Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633

Guth A. H., 1981, Phys. Rev. D, 23, 347

Haardt F., Madau P., 2012, ApJ, 746, 125

Hariharan N., Graziani L., Ciardi B., Miniati F., Bungartz H. J., 2017, MNRAS, 467, 2458

Hassan S., Davé R., Finlator K., Santos M. G., 2016, MNRAS, 457, 1550

Hassan S., Davé R., Finlator K., Santos M. G., 2017, MNRAS, 468, 122

Hayes J. C., Norman M. L., 2003, ApJS, 147, 197

He C.-C., Ricotti M., Geen S., 2020, MNRAS, 492, 4858

Hinton G. E., Srivastava N., Krizhevsky A., Sutskever I., Salakhutdinov R. R., 2012, arXiv e-prints, p. arXiv:1207.0580

Hoag A., et al., 2019, ApJ, 878, 12

Hoyle F., 1948, MNRAS, 108, 372

Huang K.-W., Di Matteo T., Bhowmick A. K., Feng Y., Ma C.-P., 2018, MNRAS, 478, 5063

Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27

Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725

Izotov Y. I., Schaerer D., Worseck G., Guseva N. G., Thuan T. X., Verhamme A., Orlitová I., Fricke K. J., 2018, MNRAS, 474, 4514

Jana R., Nath B. B., 2018, MNRAS, 479, 153

Jelić V., et al., 2010, MNRAS, 402, 2279

Kakiichi K., Gronke M., 2019, arXiv e-prints, p. arXiv:1905.02480

Kakiichi K., Graziani L., Ciardi B., Meiksin A., Compostella M., Eide M. B., Zaroubi S., 2017, MNRAS, 468, 3718

Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, MNRAS, 485, 117

Katz H., Kimm T., Haehnelt M., Sijacki D., Rosdahl J., Blaizot J., 2018, MNRAS, 478, 4986

Katz H., et al., 2020, MNRAS, 494, 2200

Keating L. C., Puchwein E., Haehnelt M. G., 2018, MNRAS, 477, 5501

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Khullar S., Ma Q., Busch P., Ciardi B., Eide M. B., Kakiichi K., 2020, MNRAS, 497, 572

Kimm T., Blaizot J., Garel T., Michel-Dansac L., Katz H., Rosdahl J., Verhamme A., Haehnelt M., 2019, MNRAS, 486, 2215

Kitayama T., Yoshida N., Susa H., Umemura M., 2004, ApJ, 613, 631

Komatsu E., et al., 2011, ApJS, 192, 18

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Krawczyk C. M., Richards G. T., Mehta S. S., Vogeley M. S., Gallagher S. C., Leighly K. M., Ross N. P., Schneider D. P., 2013, ApJS, 206, 4

Kulkarni G., Worseck G., Hennawi J. F., 2019, MNRAS, 488, 1035

Leite N., Evoli C., D'Angelo M., Ciardi B., Sigl G., Ferrara A., 2017, MNRAS, 469, 416

Linde A. D., 1982, Physics Letters B, 108, 389

Liu H., Slatyer T. R., Zavala J., 2016, Phys. Rev. D, 94, 063507

Ma Q., Helgason K., Komatsu E., Ciardi B., Ferrara A., 2018a, MNRAS, 476, 4025

Ma Q., Ciardi B., Eide M. B., Helgason K., 2018b, MNRAS, 480, 26

Ma Q.-B., Ciardi B., Kakiichi K., Zaroubi S., Zhi Q.-J., Busch P., 2020, ApJ, 888, 112

Madau P., 2017, ApJ, 851, 50

Madau P., Fragos T., 2017, ApJ, 840, 39

Madau P., Haardt F., 2015, ApJ, 813, L8

Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429

Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648

Maselli A., Ferrara A., Ciardi B., 2003, MNRAS, 345, 379

Maselli A., Ciardi B., Kanekar A., 2009, MNRAS, 393, 171

Mason C. A., et al., 2019a, MNRAS, 485, 3947

Mason C. A., Naidu R. P., Tacchella S., Leja J., 2019b, MNRAS, 489, 2669

Matsuoka Y., et al., 2018, ApJ, 869, 150

Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, MNRAS, 451, 400

Matthee J., Sobral D., Best P., Khostovan A. A., Oteo I., Bouwens R., Röttgering H., 2017, MNRAS, 465, 3637

Matthee J., Sobral D., Gronke M., Paulino-Afonso A., Stefanon M., Röttgering H., 2018, A&A, 619, A136

McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499

McQuinn M., Lidz A., Zaldarriaga M., Hernquist L., Hopkins P. F., Dutta S., Faucher-Giguère C.-A., 2009, ApJ, 694, 842

Meiksin A. A., 2009, Reviews of Modern Physics, 81, 1405

Meiksin A., Khochfar S., Paardekooper J.-P., Dalla Vecchia C., Kohn S., 2017, MNRAS, 471, 3632

Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006a, New Astron., 11, 374

Mellema G., Iliev I. T., Pen U.-L., Shapiro P. R., 2006b, MNRAS, 372, 679

Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955

Mineo S., Gilfanov M., Sunyaev R., 2012a, MNRAS, 419, 2095

Mineo S., Gilfanov M., Sunyaev R., 2012b, MNRAS, 426, 1870

Mirabel I. F., Dijkstra M., Laurent P., Loeb A., Pritchard J. R., 2011, A&A, 528, A149

Moriwaki K., Yoshida N., Eide M. B., Ciardi B., 2019, MNRAS, 489, 2471

Naidu R. P., Forrest B., Oesch P. A., Tran K.-V. H., Holden B. P., 2018, MNRAS, 478, 791

Nasir F., Bolton J. S., Becker G. D., 2016, MNRAS, 463, 2335

Nath B. B., Biermann P. L., 1993, MNRAS, 265, 241

Nelson D., et al., 2018, MNRAS, 475, 624

Norman M. L., Paschos P., Abel T., 1998, Mem. Soc. Astron. Italiana, 69, 455

O'Shea B. W., Wise J. H., Xu H., Norman M. L., 2015, ApJ, 807, L12

Ocvirk P., et al., 2016, MNRAS, 463, 1462

Oh S. P., 2001, ApJ, 553, 499

Onoue M., et al., 2017, ApJ, 847, L15

Ostriker J. P., Gnedin N. Y., 1996, ApJ, 472, L63

Ouchi M., et al., 2010, ApJ, 723, 869

Ouchi M., et al., 2018, PASJ, 70, S13

Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2015, MNRAS, 451, 2544

Paciga G., et al., 2011, MNRAS, 413, 1174

Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, MNRAS, 443, 678

Parsa S., Dunlop J. S., McLure R. J., 2018, MNRAS, 474, 2904

Parsons A. R., et al., 2010, AJ, 139, 1468

Parsons A. R., et al., 2014, ApJ, 788, 106

Patil A. H., et al., 2017, ApJ, 838, 65

Pawlik A. H., Rahmati A., Schaye J., Jeon M., Dalla Vecchia C., 2017, MNRAS, 466, 960

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Penzias A. A., Wilson R. W., 1969, ApJ, 156, 799

Planck Collaboration et al., 2016, A&A, 596, A108

Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209

Pober J. C., et al., 2015, ApJ, 809, 62

Price L. C., Trac H., Cen R., 2016, arXiv e-prints, p. arXiv:1605.03970

Puchwein E., Pfrommer C., Springel V., Broderick A. E., Chang P., 2012, MNRAS, 423, 149

Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, MNRAS, 485, 47

Rees M. J., Setti G., 1970, A&A, 8, 410

Regan J. A., Haehnelt M. G., 2009, MNRAS, 393, 858

Ricotti M., Ostriker J. P., 2004, MNRAS, 352, 547

Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19

Rosdahl J., et al., 2018, MNRAS, 479, 994

Ross H. E., Dixon K. L., Iliev I. T., Mellema G., 2017, MNRAS, 468, 3785

Ross H. E., Dixon K. L., Ghara R., Iliev I. T., Mellema G., 2019, MNRAS, 487, 1101

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471

Safarzadeh M., Scannapieco E., 2016, ApJ, 832, L9

Santos M. G., Ferramacho L., Silva M. B., Amblard A., Cooray A., 2010, MNRAS, 406, 2421

Sazonov S. Y., Khabibullin I. I., 2017, Astronomy Letters, 43, 211

Sazonov S., Sunyaev R., 2015, MNRAS, 454, 3464

Schaepman M., 2009, in Warner T., Duane Nellis M., Foody G., eds, , The SAGE Handbook of Remote Sensing. Sage, London, pp 166–178

Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, ApJ, 795, 20

Schneider D. P., Schmidt M., Gunn J. E., 1991, AJ, 102, 837

Schroeder J., Mesinger A., Haiman Z., 2013, MNRAS, 428, 3058

Semelin B., Eames E., Bolgar F., Caillat M., 2017, MNRAS, 472, 4508

Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33

Shull J. M., van Steenberg M. E., 1985, ApJ, 298, 268

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Singh S., et al., 2017, ApJ, 845, L12

Singh S., et al., 2018, ApJ, 858, 54

So G. C., Norman M. L., Reynolds D. R., Wise J. H., 2014, ApJ, 789, 149

Sokolowski M., et al., 2015, Publ. Astron. Soc. Australia, 32, e004

Springel V., 2005, MNRAS, 364, 1105

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Starobinsky A. A., 1980, Physics Letters B, 91, 99

Steidel C. C., Bogosavljević M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, ApJ, 869, 123

Strickland D. K., Stevens I. R., 2000, MNRAS, 314, 511

Stücker J., Busch P., White S. D. M., 2018, MNRAS, 477, 3230

Suchkov A. A., Balsara D. S., Heckman T. M., Leitherer C., 1994, ApJ, 430, 511

Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173

Swarup G., 1991, in Cornwell T. J., Perley R. A., eds, Astronomical Society of the Pacific Conference Series Vol. 19, IAU Colloq. 131: Radio Interferometry. Theory, Techniques, and Applications. pp 376–380

Tingay S. J., et al., 2013, Publ. Astron. Soc. Australia, 30, e007

Trebitsch M., Blaizot J., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 470, 224

VanderPlas J., Connolly A. J., Ivezic Z., Gray A., 2012, in Proceedings of Conference on Intelligent Data Understanding (CIDU. pp 47–54 (arXiv:1411.5039), doi:10.1109/CIDU.2012.6382200

Vanzella E., et al., 2016, ApJ, 825, 41

Vanzella E., et al., 2018, MNRAS, 476, L15

Vasiliev E. O., Ryabova M. V., Shchekinov Y. A., Sethi S. K., 2018, Astrophysical Bulletin, 73, 401

Volonteri M., Bellovary J., 2012, Reports on Progress in Physics, 75, 124901

Volonteri M., Lodato G., Natarajan P., 2008, MNRAS, 383, 1079

Voytek T. C., Natarajan A., Jáuregui García J. M., Peterson J. B., López-Cruz O., 2014, ApJ, 782, L9

Vrbanec D., Ciardi B., Jelić V., Jensen H., Iliev I. T., Mellema G., Zaroubi S., 2020, MNRAS, 492, 4952

Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, Nature, 585, 39

Weinberger R., et al., 2018, MNRAS, 479, 4056

Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, ApJ, 825, 144

Worseck G., Davies F. B., Hennawi J. F., Prochaska J. X., 2019, ApJ, 875, 111

Yoshida N., Oh S. P., Kitayama T., Hernquist L., 2007, ApJ, 663, 687

Zheng Z.-Y., et al., 2017, ApJ, 842, L22

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

van Haarlem M. P., et al., 2013, A&A, 556, A2

van de Hulst H. C., Raimond E., van Woerden H., 1957, Bull. Astron. Inst. Netherlands, 14, 1

Acknowledgements

Finally, to the best part.

I wish to thank my trailblazing, skilled and wise supervisor, Benedetta Ciardi, for the momentous, exciting and slightly arduous projects. They keep bearing scientific fruits. No challenges have been too big or too small for us to solve, and I am grateful for having learned to approach them in a thorough, precise way. It has been an amazing experience.

The projects would not have been tangible and approachable without the preparations and hard work of Luca Graziani and Koki Kakiichi. Thank you for this, and for all your advice and supervision. I have learned to 'just do it'. I am also grateful for the data, suggestions and encouragement from Tiziana Di Matteo and Yu Feng, who have absolutely brilliant minds.

Thank you, Max Grönke, Rahel Frick and Mark Dijkstra for being such inspiring scientists and friends. Max, thank you for all the memorable moments. And thank you, Martin, Philipp, Jens, Chris, Qingbo and Enrico for discussions, encouragement, your keen and sharp thoughts and for knowing what it means to sail on the doctoral sea. Thank you, Aoife, Inh, Dani, Ricard, Matteo, Haakon, Malin, Durand, Dijana, Isabella and Ilkham for letting me into your lives.

To my sign language interpreters. Cecilie, Maria, Anette and Silje, thank you for your unparalleled work and your good companionship. It is a pleasure to be around you.

I have been blessed with an unforgettable stay in the US. Thank you, Piero Madau, for hosting me, and thank you, Platon, Kevin, César, Amanda, Grecco, and the rest of you slugs for welcoming me so warmly. I have also always been welcome at ITA in Oslo, thank you Sijing Shen, Håkon Dahle and Kristine Knudsen for making space for me, and Robert, Mattia and Marie for making life fun. The secretaries Rickl, Kratschmann, Veith and Depner at MPA who adeptly helped me with housing and more deserve special thanks, and thanks to the directors for accepting and supporting me.

But none of anything would have been possible without my family. A big thank you to my wonderful parents, Synøve and Thomas, for all the kind things you do that I do not always see. To my little big brother David, to Camilla and to my big little brother, Carl Christian, my great siblings. And to my nephews and niece, for distractions and reminders of the important things in life. If you ever should read this, you are delightfully weird!

And at last, to my dear friends. A big thank you, Tarek, Truls, Andrea, Martin, Charlie, Maria, Asbjørn, Petter and Rémi, your support has been welcome and surprised me so many times. And finally, thank you, Anders. You are the finest kind of star an astronomer could hope to find. You are out of this world!